# Performance Analysis in Elite Sports

Talsma, Bertus Gatze



*Publication date:*
2013

# Performance Analysis in Elite Sports

Statistical Comparison, Development, Ranking, and Selection in Speed Skating and Soccer

Bertus G. Talsma

Published by: University of Groningen, Groningen, The Netherlands

Printed by: Ipskamp Drukkers B.V., Enschede, The Netherlands

ISBN: 978-90-367-6167-3



This document was prepared using the LaTeX document preparing system.

**RIJKSUNIVERSITEIT GRONINGEN**

# Performance Analysis in Elite Sports

## Statistical Comparison, Development, Ranking, and Selection in Speed Skating and Soccer

**Proefschrift**

ter verkrijging van het doctoraat in de
Economie en Bedrijfskunde
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. E. Sterken,
in het openbaar te verdedigen op
donderdag 28 november 2013
om 12.45 uur

door

**Bertus Gatze Talsma**

geboren op 1 april 1983

te Wonseradeel

| | |
|---|---|
| Promotores: | Prof. dr. G. Sierksma<br>Prof. dr. R.H. Teunter |
| Beoordelingscommissie: | Prof. dr. R.H. Koning<br>Prof. dr. F.C.R. Spieksma<br>Prof. dr. J.A.A. van der Veen |

ISBN: 978-90-367-6167-3

# Preface

This PhD project has given me the opportunity to combine two of my interests, namely sports and statistics/mathematics. For the past seven years, I have worked (most of the times) with great pleasure and enthusiasm on this thesis, and now that the work is completed, I would like to thank all people that have supported me.

First of all, I would like to express my thankfulness to my supervisor Gerard Sierksma, who shares my interest in sports and stats/math. He suggested to me to become a PhD student, and I am proud to be his last student. Gerard: "I would like to thank you for all the help and effort that you put into this project, especially for your patience and for the countless fruitful discussions we had about the content of this thesis and relating sport issues." Furthermore, I would like to thank my second supervisor Ruud Teunter for his support during the final stages of the project.

I would also thank the members of the dissertation committee: prof.dr. R.H. Koning (University of Groningen), prof.dr. F.C.R. Spieksma (KU Leuven), and prof.dr. J.A.A. van der Veen (Nyenrode Business University) for their time and the valuable comments they gave me.

I am grateful to SOM, which provided me the support I needed in the more difficult times and the extra time they allowed me to complete this thesis. In particular I am thankful to Martin Land, Linda Toolsma, and Jet ter Heegte.

I want to thank the following people I worked with during my PhD period and who contributed to this thesis. Richard Kamps, Ronald Schoonbeek, and Egbert van der Veen are acknowledged for using their bachelor theses in support of my work. The KNSB co-operators, and especially Arie Koops, are acknowledged for the joined work regarding the Olympic speed skating selection procedures. Ronald Kres is acknowledged for collecting the soccer data, and dr. Gerard Kuper for co- writing the 1000m-paper. All colleagues of ORTEC-TSS, the company where I worked part-time during my PhD period and where I have started working after the completion of this thesis, are acknowledged for creating an inspiring work environment.

Furthermore, I would like to thank my office mates, Lolke, Marcel, Dimtri, and the fellow PhD candidates and colleagues of the Operations Department for the great time we spent together.

Not only work related but also outside work, I would like to thank the people who supported me during the research process. All members of FDTD: “Due to you I spent too many hours in Middle Earth, but it was always fun with you on TS.” My teammates from the VV Blauwhuis soccer team: “Although we did not always win, it was always a pleasure to come back from Groningen and to spend the Sunday with you on the soccer field.”

My classmates from secondary school, Klaas Hendrik, Jelte, Niels, Steven, Hendriekje, José, Sylvia and Rina: “I have always enjoyed our traditional Christmas dinners and I hope we will continue this tradition for a long time.”

My two friends and paranymphs, Richard and Arjen: “Thank you for all the time we spent together and your support during my final defense.” Arjen: “It was always a pleasure to spend my Saturdays with you while working outdoors, clearing my head during cutting and piling wood logs. But ’five o’clock’ I liked the most.” Richard: ”I greatly value our friendship. You are a true friend and I always enjoy hanging out at your place, just to relax, playing games and drinking beer. Your humor and perspective motivated me in moments when I needed it and kept me going.” “Thank you both for everything.”

Furthermore, I would like to thank my family and all of my other friends for their support.“You all motivated and encouraged me to finish my thesis.”

*As lêste wol ik heit en mem tanksizze foar al harren stipe en soargen, dy’t sy my jûn hawwe yn dizze perioade. It wie altyd noflik om wykeins wer efkes thús te wêzen. Tige tank.*

# Contents

Chapter 1

# Introduction

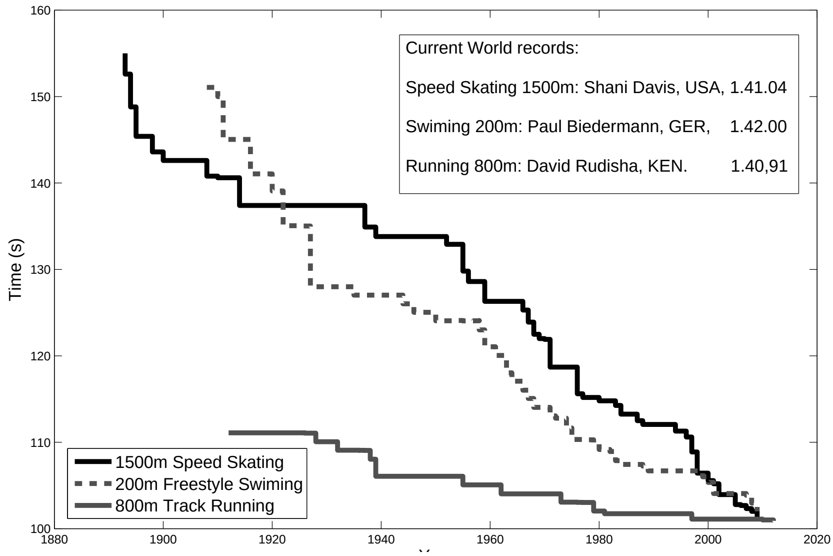
## 1.1 Introduction

**Development in Sport Performances** Becoming better than opponents and reaching the top are major goals of top sport athletes. These goals are achieved by constantly improving the performance through, among others, applying new training methods and trying out new tactics. In addition to the pure human input, i.e., the athletic ability, also 'external' aspects play a role in performance improvement and reaching the top. In particular, the implementation of technological innovations may lead to better performance.

One of the most common gauges for measuring ongoing performance improvements is the development of world records. Figure 1.1 shows the development of three world records of sport disciplines that have in 2010 more or less the same finishing time. These disciplines are the 1500m speed skating, the 200m free style swimming, and the 800m track running. The graphs in this figure illustrate the difference between a sport in which mainly the human input determines the performance (e.g., running), and sports in which also 'external' factors, such as technological innovations, play a key role (e.g., swimming and speed skating). The three graphs show that, although the three world records are currently almost equal, speed skating and swimming have improved significantly more than running.

One of the 'hidden' causes of the constantly improving performances is the "spread of excellence" effect, as it was called by Stephan Jay Gould, an American evolutionary biologist. This effect, which we will call the Gould effect (in the literature also called Gould's hypothesis; see, e.g., Schmidt and Berri (2005)), is described in Gould (1996). Gould describes in this book the internal dynamics of complex systems. In case of sports, this means a constant improvement of the performances due to just practicing the sport, see Section 2.3.4. Not only does the effect lead to a certain leveling of performances at the top, it also makes the differences in performances between top athletes smaller and smaller over time. More and more the limits of what is humanly possible are reached. In Scully (2000), these boundaries are analyzed for

**Figure 1.1.** World record development




track-and-field events.

Gould's effect can also be observed in speed skating and swimming. In these sports, the performance differences for certain events are already so small that they become hardly or even impossible to measure because they lie within the error margins of the time registration systems. A famous example of such a situation is the gold medal race of Michael Phelps during the 100m butterfly race at the Olympic Games of 2008. The debate concerning his first place in this race is still going on. Also the photograph of the finish of the 1000m race in 2007 in Calgary of the two speed skaters Shani Davis and Simon Kuiper is classic; see Figure1.2. Although the photograph leaves no doubt about the winner Kuiper, the time registration systems declared Davis as the winner. The interesting point here is that if the two skaters had competed in different runs, then this picture could not have been taken, and Davis would have been the winner; see also Korver and Hoogveld (2008). The problems with time registration in top sports are also demonstrated in the Dutch video report NOS (2012).

Part of the sport science literature concerns the analysis of performance improvements, and tries to contribute to the understanding and development of performances. Various studies report performance analysis of individual athletes and of teams; see e.g., Nevill *et al.* (2008). Hughes and Bartlett (2002) present a wide range of indicators, distinguishing between general match, tactical, technical, and biomechanical indicators. The influence of, among others, biomechanics (e.g., Yeadona and Challisa (1994) ), nutrition (e.g., Rodriguez *et al.* (2009)), techniques (e.g., Lees (2002)) and tactics (e.g., Reep *et al.* (1971) ) on the improvement of sport performance are popular areas of research. The study of team sports in this respect is divided into

**Figure 1.2.** Fotofinish, 1000m race, 2007, Calgary.

the following categories: wall and net games, like squash (Hughes *et al.* (2007)), invasion games, like basketball (e.g., Niemi (2010)) and soccer (e.g., Hughes and Franks (2005)), striking and fielding games, like baseball (e.g., Albright (1993)). Individual sports, like track-and-field (e.g., Grubb (1998)) and swimming (e.g., Trewin *et al.* (2004)) are also frequently investigated in this respect.

**Performance Comparison** A major part of the study of performance development is performance comparison. During sport events, such as matches, tournaments, and competitions, performance comparison is daily practice: without comparison there is no ranking and no winner. Of course, the final results are subject to a number of usually well-defined rules. These include not only the rules of the game, but also rules for the ranking. For some sport events, such as a marathon or a soccer match, the ranking rules are straightforward because they are based on (more or less) objective quantitative measurements: the athlete that runs the fastest, or jumps the highest is the winner. Or, in case of team sports, the team with the highest number of goals is the winner. For tournaments, the ranking rules are usually more complicated, and it is not always true that the athlete with the best performance during the tournament is the winner. It may happen that the winner of the final round of a tournament receives a gold medal without having set the fastest race time of that tournament.

Many of the rules, both game rules and ranking rules, have changed over the years. In soccer, for example, before 1994 two points were awarded for the winner of the match; in 1994, this number was increased to three points. Of course, different ranking rules may result in different rankings and standings. Besides this, also the rules of the game of practically every sport have undergone changes throughout the years.

For some sports the ranking rules are even more complicated. This holds especially for multi-event sports, such as the decathlon in track-and-field. Here the results on the various events are measured and aggregated into one score. Due to the complicated final ranking rules it may be very hard to balance the training for the various decathlon events; see Zwols and Sierksma (2009).

Final results of matches and tournaments do not always reflect the right performance comparison. The results may be influenced by wrong referee decisions and unobserved breaches of the rules. Especially in soccer, where the number of goals is small and the financial interests are high, nowadays, wrong referee decisions may have completely undesirable and far-reaching consequences for the losing teams.

The use of doping also biases performance comparison and is, especially after the USADA report (USADA (2012)) concerning cyclist Lance Armstrong, under broad discussion nowadays. On the other hand, dope sinners are of all times. It is assumed that cyclist Arthur Linton was one of the first doping users: he died of an overdoses trimethyl. In 1928, the International Association of Athletics Federation (IAAF), the governing body for track-and-field, became the first international sports federation to prohibit usage of doping by their athletes, see International Association of Athletics Federation (2009). Yet each year offenders are caught, in practically all sports disciplines.

Performance comparisons across different events are sometimes necessary when for example athletes need to be selected. However, the selection is usually based on results obtained under different circumstances and at different time instances. Also when recent performances are compared with those from years ago, the performance results need to be corrected for changing circumstances, such as different competition levels, different ranking and tournament rules, and in many cases different technological support. Also the process of maturing of the sport (Gould's effect) has to be taken into account.

**Comparison and Ranking** Performance comparison within single sport events is more or less straightforward. The results during the event directly determine the ranking of the athletic performances. For performance comparison of multiple events or performance comparisons across different time eras, the rules are not always clear, and comparisons are complicated.

Sport performance rankings can be divided into, what we call, static and dynamic rankings. Static rankings order the athlete's performances in a sport event and, at the end of the event, the ranking determines a winner. Dynamic rankings have, in contrast to static rankings, no end date and change constantly after each event. We may distinguish the following types of static rankings:

- Single event rankings;
- Integrated single event rankings;
- League and tournament rankings;
- Multi-competition and multi-tournament rankings.

Single events are sport events that stand on their own. The final ranking is completely determined by that event. Examples of single events are the marathon, the Super Cup in soccer, and the Dutch 11-city race for speed skating. In contrast to single events, integrated single events have a winner but are part of a league or tournament for which there is an overall winner. Examples are the individual distance races during allround speed skating tournaments, and the Tour de France stages.

League and tournament rankings are calculated from the results of the different stages and are usually aggregated scores of the individual integrated single events. Multi-competition and multi-tournament rankings combine the results of several individual leagues or tournaments. Examples include the UCI ranking in cycling, the World Cup competition in speed skating, and the Olympic Games medal rankings of countries. In most cases a point-based system is used: points are awarded to the ranking positions on the various competitions or tournaments that are taken into account.

Examples of dynamic rankings are the ATP ranking in tennis, the FIFA world ranking in soccer, the Adelskalender in speed skating, and all world record rankings. Some of these dynamic rankings have a shifting start date, like the ATP and the FIFA world ranking, and rank over a fixed period of time. This means that for these rankings, athletes appear and disappear continuously. Other dynamic rankings have a fixed starting date and contain a fixed selection of athletes of the corresponding sport discipline, like the Adelskalender.

A major problem when comparing performance in dynamic 'all time' rankings that span a few decades is that they are usually heavily biased by the fact that the concerning performances may have been delivered under very diverging circumstances. Also not all athletes that are included have competed directly against each other. In order to make a 'fair' comparison, new rules and assumptions need to be formulated and validated. Furthermore the results need to be corrected for the diverging circumstances. For example, in Chapter 2, we use statistical techniques to correct speed skating results, leading to a situation that all athletes perform under more-or-less identical circumstances. So also the athletes of the 'old' days have a fair chance to reach the top of the recent rankings.

A widely applied technique for ranking sports performances is Data Envelopment Analysis, DEA. DEA is a non-parametric linear optimization-based technique for measuring the relative efficiency of a set of decision making units (DMU's) which consume multiple inputs to produce multiple outputs; see, e.g., Soleimani-Damaneh *et al.* (2011). DEA models use teams or athletes as DMU's and the rankings are based on an efficiency relation between input and output. The more efficient the DMU is, the more it distinguishes itself from the other DMU's. For an elaborate treatment of DEA models as ranking models, we refer to Adler *et al.* (2002).

In Estellita-Lins *et al.* (2003) and Wu *et al.* (2009), the Olympic country classification based on the number of medals is calculated by measuring the efficiency of countries taking into account population and the gross domestic product. In Haas (2003) and Bosc *et al.* (2009) the performances of soccer teams are compared using DEA models. In Haas (2003) the efficiency of teams is based on points and revenues using as input players and coaches' salaries. In Bosc *et al.* (2009) the efficiency of the offensive and the defensive lines of the teams are analyzed by applying DEA models.

In this thesis we do not use DEA models, because we are mainly interested in identifying the best performances of teams and athletes, and not so much in the efficiency of team and athlete performances. For the specific questions of this thesis we are not so much interested in the way results are obtained, but more in the final results and their qualities in relation to the performances of the competitors.

**Comparison and Selection** Also in sports, selection takes place at several levels. Trainers select players for their line-up, federations select athletes to participate in tournaments (e.g., the Olympic Games), or they select young talents to continue their careers as professionals. Selection procedures are often the cause of intensive discussions. Not being selected could mean not participating in a major tournament or not becoming a professional athlete; see Williams and Reilly (2000). This means that stakes are high. Of course whether or not an athlete is selected is based on his/her expected performance, but the estimation of an expected performance is often disputable. On the other hand, past performances are most of the times the only objective information on which an expected performance can be based. However, a major problem of comparing past performances is that they may be delivered under different circumstances. In order to compare the performances, again corrections need to be made on these differences.

In 2012, the Kenyan athletic federation had to choose their three marathon runners for the Olympic Games in London. At that time, the country counted 278 marathon runners with the A-status. In comparison, only 43 European runners satisfied this criterion. The first twenty positions on the world ranking of fastest marathon runners were occupied by Kenyan athletes. The athletic federation started the selection procedure with a pre-selection of six runners made by the federation itself; the final decision was based on their performances in April 2012, about four months before the Games. Of these six runners, four ran the marathon in London, one in Rotterdam and one in Boston. Of course, the circumstances during these marathons were different. So, how to make a fair comparison between the results?

During the Olympics the gold medal went to Kiprotich from Uganda, leaving Kenya with position two and three. The question whether or not the federation had selected the right athletes will never be answered. One way to deal with the comparison problem is to arrange a trial, like the USA usually do. On the other hand, it would also have been possible to correct the race times of all Kenyan athletes for the various circumstances.

Although high expected performances are no guarantee for success and projections may differ from realizations, they are a useful tool for the analysis of the realizations. Expected performance can be used as benchmarks of the real competition results. Question like: Why has the athlete performed so much better or worse than expected? can then be answered more accurate and effectively.

A drawback of selection tournaments is that athletes may focus on such events, resulting in a disturbance of the periodization concerning the preparation for the final major tournament. Estimations based on selection moments and previous results therefore do not always provide the most relevant information. Where some athletes are at their top during the selection moments just for becoming selected, other athletes are still training, building up their performances such so to peak at the final major tournament.

**Fairness** Fairness and sport is a combination that sounds like a *contradictio in terminis*. Especially in competitive sports, athletes try to distinguish themselves by better training, better nutrition, better equipment, and sometimes 'better' rule management. The main goal usually is to beat the opponent. The moral rule of 'maximize rule following', in other words strictly follow the rules, is more and more exchanged for its 'dual' rule of 'minimize the probability of being caught'. Of course, all actions within the rules are fair, but in reality 'fair play' means 'not being caught by the referee'. This is particularly true for the usage of doping. Doping sinners can be prosecuted even eight years after the violation, even if the violation is not established during the match or tournament. Thus, in the case of doping, 'fair play' means 'not being caught during the game or in the eight years following the game'.

A generally accepted condition of sport events is that all participants compete under equal circumstances. Obviously, this cannot always be guaranteed. Sometimes the weather conditions change during an important event. Sometimes, the circumstances are different by rule: a skater who starts in the inner lane during a 1000m race traverses a different route than the skater who starts in the outer track. While this seems like a rather minor difference, we show in Chapter 5 that the difference is actually significant for the women. Also 'home advantage' can be considered as an intentional unfairness and widely studied across several sports; see the overview paper of Nevill and Holder (1999).

So, although performances are supposedly compared under equal circumstances, this is very often not the case, and one has to consider whether or not the circumstances create significant benefits. In case the different circumstances are seen as unequal or create a significant advantage, one has to examine the rules of the game and see if they can be changed in order to eliminate the unfairness.

Comparing performances of athletes from different time eras is only 'fair' if the performances are not biased by the changing circumstances. Comparing skaters from before the klapskate period with riders that have grown up with the klapskate is only fair if the results are appropriately corrected; see Chapter 2.

Classical examples of equipment unfairness are the following cases. In 1997, Tony de Jong was the only speed skater with klapskates during the European Championships. Completely unexpectedly, she beat multiple European and world champion Gunda Niemann. The introduction of 'high tech' body suits in competitive swimming in 2008 led to heated discussions; see Partridge (2011). The new suits enabled most of the swimmers to improve their personal records, and during the Olympics of Bejijng in 2010 practically all swimmers were much faster than expected; see Brammer *et al.* (2009). Some swimmers even used several of these high tech suits on top of each other so as to increase the floating effect. In 2010, the Federal International de Natation (FINA) introduced new rules regarding the suits and disbanded them. Note that there is a subtle difference between the case of the klapskate and case of the swimming suits. Every skater had the opportunity to buy his pair of klapskates, but not all swimmers had access to the best swimming suits. This was because of sponsor contracts that prescribed the suits to be used during the race.

It should be clear from the above discussions that matches need to be organized under equal conditions for all participating athletes, and only the differences in athletic ability should determine the outcome. Obviously, this is not always guaranteed in practice. Especially the fact that the mutual differences at the top are becoming smaller and smaller, measuring the length of the tracks and measuring the starting and finishing times, together with offering equal tracks, at least to the potential gold medal winners, becomes crucial for deciding the right winners. Using error margins for the final results may partly solve these problems, in the sense that for example a new world record should be outside the error margin of the old record.

**Overview** In chapter 2, the performance development of speed skaters is analyzed and all time dynamic rankings, which we call Universal Speed Skating Rankings, are presented. Already for more than one hundred years speed skaters have been competing against each other in international competitions and tournaments. We have collected and used all race results from the period 1892-2010. Due to changing circumstances, the uncorrected race results cannot be used for a 'fair' ranking. The current generation of riders skate a lot faster than their colleagues from the old days, and this is for a large part not due to the fact that the current generation consists of more skilled skaters. This chapter presents a comparison method by formulating, among others, so-called pre-conditions for the final rankings. These pre-conditions serve as a certain format within which the actual ranking rules are formulated and the results are validated.

The first part of Chapter 2 discusses the changing circumstances and how they have influenced the actual performances. Rink and equipment improvements have caused performance jumps, whereas a continuous improvements can be related to Gould's effect and the thereby related changing skating population. The second aspect of Gould's effect, concerning the decreasing variation between performances, is made clear from the data. In order to make a fair comparison of performances, we introduce a special performance measure that corrects for most external factors. So, only the performance dimension concerning the human output remains. In fact, these dimension values are used for the rankings. Finally, existing ranking models are tested and compared against the, already mentioned, pre-conditions.

In the second part of the chapter the ranking system is presented, and the ranking lists are calculated. We have designed all time dynamic rankings for allround and sprint speed skating for both for men and women. In addition, we present lists for all distances. For validation purposes, extensive scenario analyses are applied.

Chapter 3 deals with the performance comparison of professional Dutch soccer players. This Chapter is based on Schoonbeek (2010). Whereas speed skating is an individual sport, soccer is a team sport. Therefore different comparison tools are required. An important assumption that we use is the fact that each team player's contribution to the match result is proportional to the number of minutes the player was active during that match. We have, among others, corrected for the fact that attackers/strikers have on average less playing minutes than goalkeepers. Also the importance of the match is taken into account: final matches are rewarded considerable higher than first round matches. Since the introduction of professional soccer in 1954, Dutch players have been active in all international tournaments. We have restricted the comparison to the international competitions, because acting in the international arena provides the best standard for comparing performances.

Chapter 4 is devoted again to speed skating. Now, the performances of individual speed skaters are compared in order to select skaters with a highest probability to win medals on the Olympic Winter Games. The difficulty here is that performances of different skaters on different distances need to be compared. This is due to the fact that, both for women and men, there are eighteen starting positions on five distances, but only ten skaters with the same nationality can be selected according to the IOC rules. This chapter describes an assignment model of which the coefficients of the objective function are the various winning probabilities. The second part of the chapter describes the actual selection procedure. This selection procedure has been implemented and used by the Dutch Speed Skating Association (KNSB) with respect to the Olympic Winter Games of 2010. We also describe the role of the performance matrix (prestatiematrix) in the KNSB selection procedure.

In Chapter 5, we analyze the concept of fairness in case of the 1000m race in speed skating. This chapter is published as Kamst *et al.* (2012). In every speed skating race, one skater starts in the inner lane and another one in the outer lane. For the 1000m speed skating this means that one rider skates three inner curves and two outer, whereas the other one skates two outer and three inner curves. If the medals are distributed based on the results of only one race, as happens during Olympic Games and World Single Distances Championships, this difference may be qualified as a kind of unfairness. Already in 1994, Hjort (2004) discovered a significant difference between the 500m inner lane times and the outer lane times. Based on these findings the ISU decided to always skate the 500m twice, making sure that each skater has one start in the inner lane and one in the outer lane. In Kamst *et al.* (2010), Hjort's analysis is repeated for the data from the period 1997-2010, the period after the introduction of the klapskate. Surprisingly, no significant differences occur anymore. In the case of the 1000m, the results are different. For the women's races we found a significant differences. These differences are not significant for the men's races. Taking into account this difference, together with the above described lane differences, the fact that riders start at different positions, plus the fact that the differences at the top have become very small, we have advised the ISU to change the 1000m rules and let athletes skate every 1000m race twice.

All chapters can be read separately, and have their own conclusions and agendas for further research.

Chapter 2

# All Time Performance Comparison: Speed Skating Rankings

## 2.1 Introduction

### 2.1.1 Ranking systems in sports

People are always eager to know who is better, richer, or smarter. In the world of sports questions like “Who is the best?” or “Who is the best of all times?” are major questions. During competitions or events, ranking lists are made to indicate how well athletes or teams have performed, and to determine the winners. Most of these ranking systems are simple. In team sports, for example, points are awarded for victories and draws, and the team with the most points at the end of the competition is the winner. Ranking systems for most individual sports, like track-and-field and swimming, compare results during one event: the athlete with the fastest or highest score is the winner.

Next to these common rankings there exist more complicated systems. These systems summarize performances of athletes or teams across several competitions or tournaments in a certain period of time. Well-known ranking lists of this type are the ATP-ranking for tennis, the FIFA world ranking for soccer, and the UCI-ranking for cycling [1]. In Stefani (1997) the calculation systems of the first two systems are discussed; the UCI-ranking is explained in UCI (2011). Another familiar ranking system, often applied during Olympic Games, is the hierarchical or lexicographical ranking system. This system ranks countries during the Olympics on the number of medals gained; see Wittkowski (2004).

Most of these methods rank and compare only active athletes or teams. Things

[1] ATP: Association of Tennis Professionals; FIFA: Federation Internationale Football Association; UCI: Union Cycliste Internationale

become more complicated when performances from different eras are compared. In Berry *et al.* (1999) a model is presented, which compares and ranks golf, baseball, and ice hockey players from different time periods. By taking into account the potential of the players, the effects of aging, and the relative difficulty of each period, Berry *et al.* (1999) determine the best player of all times for these three sports.

### 2.1.2 Adelskalender

The most well-known all times ranking system for speed skating is the Adelskalender. The Adelskalender ranks skaters on their best performances on the four classical disciplines. However, due to technological developments and innovations none of the old champions can be found within the first hundred of the list. For example, one of the greatest skaters of all times, Eric Heiden, currently possesses position 369 on the Adelskalender. When Heiden won his fifth gold medal during the Olympics of 1980, the well known Dutch sports reporter Mart Smeets yelled: “This time will never be improved”. But now, more than twenty years later, most professional skaters are faster and are much higher ranked on the Adelskalender.

So, as the Adelskalender does not answer the question of who is the best skater of all times, how else should we compare skaters? This question was also raised by the International Skate Union (ISU) president Ottavio Cinquanta. In Snoep (2004) he introduced the idea of splitting the time scale into periods and to make separated rankings for each period. By using different rankings for different periods “the old champions will get the honor they deserve”, he suggested.

### 2.1.3 Types of ranking lists

The suggestion of Cinquanta, mentioned in Section 2.1.2, is not used in our analysis. Instead of making separate lists, we present a system that allows the inclusion of the complete time scale into one ranking. Since speed skating knows seven disciplines, both for men and women, fourteen ranking lists are designed, namely, five lists for the individual distances (men: 500m, 1000m, 1500m, 5000m, and 10000m; women: 500m, 1000m, 1500m, 3000m, and 5000m), two sprint lists (combination of 500m and 1000m), and two overall list containing all distances. Actually, for both men and women, the latter lists will give an answer to the question of who is the best skater of all times.

### 2.1.4 Ranking speed skaters

Our main objective is to rank all skaters of all times. The first problem concerns the jump-wise decreasing skating times, mainly caused by the introduction of new technologies, such as tight fit suits, klapskates, and indoor rinks. In Section 2.3.2, a survey is presented of all innovations that took place in the history of speed skating. The influence of these innovations will also be discussed there.

A second complication is the introduction of new tournaments by the ISU: new titles could be won and more races per season where skated. How to compare these results to the results of seasons in which these tournaments where not held? The introduction of new tournaments, especially the World Cups, also increase of the number of skaters participating on the highest level. In Section 2.3.5, we analyze the development of the number of participants at tournaments and the total skater's population, and discuss how it may have influenced the performances of the skaters.

The introduction of new tournaments has also caused a partition in the speed skating population. In the past, skaters were 'allround' skaters, meaning that they competed on all distances (see Section 2.2.2). Since the introduction of the single distance tournaments, more and more skaters began to specialize in one discipline or distance. This divided the skaters into either sprint or long distance specialists. The effects of this development will be discussed in Section 2.3.3.

### 2.1.5 Preview of this chapter

Section 2.2 explains the basics of speed skating. In Section 2.2.4 the data set that will be used for ranking is given. The above described problems such as the influence of innovations, the introduction of new tournaments, and the population changes are discussed in Section 2.3.1. In Section 2.4 criteria for the ranking system are formulated. In Section 2.5 we present existing rankings based on victories and world records. In Section 2.6 it is described how skating times are transformed into a new performance measure which makes skating times mutually comparable. Section 4.8 provides a ranking model that uses this performance measure. Results and validations of the model are presented in Section 2.8. In Section 2.9, sensitivity analysis on a number of input parameters of the model is carried out. Section 2.10 contains the conclusions.

## 2.2 Speed skating: basics and rules

Already in the middle ages, people used bones to skate on ice. In 1763, the first official speed skating race was organized on The Fens in England, organized by the National Ice Skating Association Bird (2001). At the end of the nineteenth century also informal international speed skating competitions were organized, which resulted in 1893 in the foundation of the International Skating Union (ISU) by fifteen countries (see ISU (2010)). With the establishment of the ISU, firm rules were made and a foundation for international competitions in speed skating was laid down. Since 1893, the ISU supervises most international skating tournaments and formulates all rules and regulations. In the following section the rules of speed skating, the various distances and tournaments, and the time and score registration methods are discussed.

### 2.2.1 Speed skating rink

In 1893, the ISU established that all international tournaments are to be organized on a 400 meter oval (see Figure 2.1). A 400m oval is divided into two separate lanes, an inside and an outside lane. The lanes have a minimum width of four meters, and the 180 degrees curves have a radius between 25 and 26 meters.

In Figure 2.1, the start and finish positions of all distances are indicated by a vertical line on the lanes. At the start of a race, the two skaters are in separate lanes, and after each lap they switch lanes. The dotted line in Figure 2.1 shows the track of a skater who starts in the inner lane at the 10000m. At the top of Figure 2.1 the skater changes lanes and after one lap he passes the start line in the other lane. After skating the full distance, the time is recorded at the finish line.

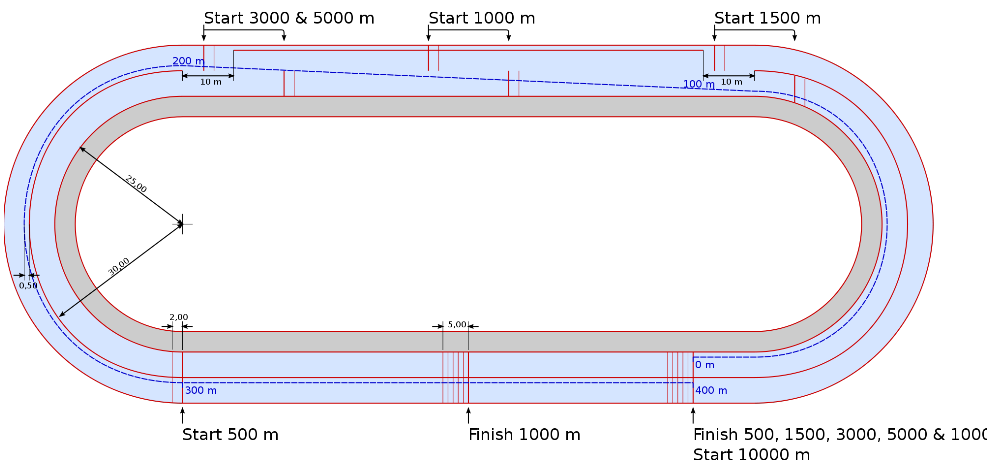



**Figure 2.1.** 400m speed skating track

### 2.2.2 Distances and tournaments

The ISU supports international races on the distances 500, 1000, 1500 meters (called short distances), and 3000, 5000, 10000 meters (called long distances). These distances are skated in various combinations at allround tournaments, sprint tournaments, and single distance tournaments. Currently, there are seven international tournaments organized by the ISU, namely, the World Allround Championships, the European Championships, speed skating at the Olympic Winter Games, the World Sprint Championships, the World Cup Competition, and the World Single Distances Championships.

The first official World Allround Championships for men under the supervision of the ISU was held in 1893 in Amsterdam. Later that year, also the first European Championships were organized. Since then, the allround championships for men have been organized annually, with the exception of the world war periods 1915 - 1921 and 1940 - 1946 (see also Table 2.1). Until now, the distance composition of allround tournaments for men has not been changed. During one weekend the athletes have to skate four distances, namely the 500m, 1500m, 5000m, and 10000m. The method for determining the winner has changed somewhat over the years. Until 1928, a skater who won three of the four distances was declared as the winner, otherwise no winner was declared. In 1928 the so-called *samalog system* (explained in Section 2.2.3) was introduced and this system is still in use.

The first World Allround Championships for women were organized in Stockholm in 1936 and since then held every year, with the exception of the period 1940 - 1947. For the women tournaments however, the distances have changed somewhat over the years. In the period 1936 - 1955, the 500m, 1000m, 3000m, and 5000m were skated. In 1956, the 5000m was replaced by the 1500m, while in 1983 the 1000m was removed to make place for the 5000m again. The same changes took place in the European Championships which were held for the first time in 1970. Due to the lack of interest the tournament was not organized between 1975 till 1981.

Speed skating events were also present at the first Olympic Winter Games in 1924. From that year on, every four years, the male skaters could win Olympic medals on the individual distances: 500m, 1500m, 5000m, and 10000m. Only at the first edition of the Olympics also medals were rewarded for the allround ranking. At the Olympic Winter Games of 1960, women made their first appearance. They competed on the 500m, 1000m, 1500m, and 3000m. In 1976 the 1000m for men was included and in 1988 the 5000m for women.

During 1973 and 1974, a small group of male skaters tried to become professional skaters and started, besides the ISU tournaments, their own world and European tournaments. After two years these tournaments disappeared again, due to bankruptcy of the organization (see Dalby *et al.* (2006)).

In 1972, the first World Sprint Championships, both for men and women, were organized in Sweden. This tournament was specifically organized for sprint specialists. In two days the sprinters have to compete on the 500m and the 1000m, each day both distances once. Before the introduction of this sprint tournament, sprinters competed along during allround tournaments and usually skipped the long distances.

To increase the number of contests, the ISU introduced the so called World Cup Competition. A competition in which the skaters competed four till eight times during the season on only single distances. The first official World Cup took place in the season 1986/1987. At the end of the season, a total of eight World Cup titles are awarded, namely for men 500m, 1000m, 1500m and 5000m/10000m (combined), and for women 500m, 1000m, 1500m and 3000m/5000m (combined).

In 1996, the ISU decided to organize world championships per single distance, the World Single Distance Championships; they are organized each season, except for the Olympic years. This tournament gives specialist the opportunity to win world titles on single distances: men on the 500m, 1000m, 1500m, 5000m, and 10000m, and women on the 500m, 1000m, 1500m, 3000m, and 5000m.

In Table 2.1 we have listed all speed skating tournaments. The first column contains the official names, the second column the acronyms of these names, the third column indicates whether the distances of the championship are rewarded individually (Sin) or combined (Com) to one overall score, and in the fourth column the periods in which the tournaments were organized are listed.

**Table 2.1.** Tournaments of speed skating

| **Men** | | | |
|---|---|---|---|
| **Tournament name** | **Acronym** | **Ind/Com** | **Seasons of organization** |
| World Allround Championships | WACh | Com | 1892-1914,1922-1940,1946-now |
| European Championships | ECh | Com | 1892-1914,1922-1940,1946-now |
| World Sprint Championships | WSCh | Com | 1970-now |
| Olympic Winter Games | OG | Ind | 1924, 1928,...,1992 (ex,1940 1944), 1994, 1998,...,2010 |
| World Cup Competition | WCC | Ind | 1979, 1986-now |
| World Single Distances Championships | WSDCh | Ind | 1996-2001,2003-2005 2006-2009,2011-now |
| **Women** | | | |
| World Allround Championships | WACh | Com | 1936-1940,1947-now |
| European Championships | ECh | Com | 1970-1974,1981-now |
| World Sprint Championships | WSCh | Com | 1970-now |
| Olympic Winter Games | OG | Ind | 1960, 1964,...,1992, 1994, 1998,...,2010 |
| World Cup Competition | WCC | Ind | 1979, 1986-now |
| World Single Distances Championships | WSDCh | Ind | 1996-2001,2003-2005 2006-2009,2011-now |

Sin=score per single distance, com = Combined score of all distances.

### 2.2.3 Winning scores, the samalog system

A crucial aspect of speed skating is the time measuring. In the early days the times were clocked by hand; nowadays electronic devices are used. The time starts running when the referee gives the start signal by firing a starting gun. During the race, lap times are recorded, and the final time is measured when the tip of the skate passes the finish line.

For championships, where single distances are skated (see Table 2.1), the skater with the fastest time on that distance is declared as the winner, with the exception of the 500m. In Hjort (2004), it is shown that skaters, starting in the inner lane at the 500m, had an advantage of 0.05 seconds on average. Based on this findings, the ISU decided to let skaters skate the 500m twice during Olympic Games and World Single Distances Championships, starting once in the inner and once in the outer lane. For the allround championships the time-based *samalog system* is used. In this system, the times of the four distances are converted to '500m times' and then added. For instance, the 1500m time is divided by 3, the 5000m by 10, and the 10000m by 20. The sum of these four '500m times' yields a point total. The skater with the lowest point total is the winner. The World Sprint Championships use the same method. Here the converted '500m times' of the two 500 meters and the two 1000 meters are added. Again the winner is the one with the lowest point total. Table 2.1 summarizes the various speed skating tournaments, together with the seasons of organization.

Next to victories on championships and single distance events, skaters have the opportunity to skate world records. A world record time is the fastest time ever skated on that distance during an ISU recognized tournament. New world records during championships do not yield extra points to the championship score, although a world record will give the skater extra status and sometimes a financial bonus.

### 2.2.4 Data set

This research uses the data set from Heijmans (2001), containing all skating results from the tournaments and distances listed in Section 2.2.2 through 2005. From 2006 on, we created our own data set. The complete data set will be denoted by $DS$ and is split into male times and female times. The male data set contains 61207 skating times, registered between 1893 and 2011. In these 119 years, 2338 male skaters have been active. Although in 1936 the first World Allround Championships for women were organized, the female data set starts in 1947. The first four years of women speed skating are not in the data set because in these years only few women participated, while after these four years, for a period of eight years, no tournaments were organized. For the period 1947-2011, the data set for women contains 40947 registered times of 1116 female skaters. Throughout this chapter the symbol $t$ always means the season $(t-1)-t$, running from September in year $t-1$ until March in year $t$. Both for men and women, the database has four dimensions, denoted by

$$
\begin{aligned}
DS &= (S_M \cup S_W) \times (Y_M \cup Y_W) \times K \times (D_M \cup D_W) \\
\text{with} & \\
S_M &= \text{the set of male (M) skaters;} \\
S_W &= \text{the set of female (W) skaters;} \\
Y_M &= \{1893, \ldots, 2011\}\text{, the set of seasons for men;} \\
Y_W &= \{1947, \ldots, 2011\}\text{, the set of seasons for women;} \\
K &= \{\text{OG, WACh, ECh, WSDCh, WSCh, WC}_1, \ldots, \text{WC}_8\}, \\
& \quad \text{the set of tournaments;} \\
D_M &= \{500\text{m}_I, 500\text{m}_O, 1000\text{m}_I, 1000\text{m}_O, 1500\text{m}, 5000\text{m}, 10000\text{m}\}, \\
& \quad \text{the set of distances for men;} \\
D_W &= \{500\text{m}_I, 500\text{m}_O, 1000\text{m}_I, 1000\text{m}_O, 1500\text{m}, 3000\text{m}, 5000\text{m}\}, \\
& \quad \text{the set of distances for women.}
\end{aligned}
$$

The labels $WC_1, \ldots, WC_8$ refer to the eight World Cups organized in one season. The notations $-_I$ and $-_O$ refer to a race with a start in the inner and outer lane, respectively, and is only used to distinguish between two of the same distances races during one tournament. For tournaments where either the 500m or the 1000m is skated only once, the subscript for the starting lane is not used. So $(i, t, k, d) \in DB$ means that skater $i$ has a registered time in season $t$ during tournament $k$ on distance

$d$. For each $(i, t, k, d) \in DB$, define

$$
\begin{aligned}
R_{itkd} &= \text{ranking of skater } i \text{ in season } t \text{ during tournament } k \text{ at distance } d;\\
N_{tkd} &= \text{number of participants in season } t \text{ during tournament } k\\
&\quad\ \text{at distance } d;\\
T_{itkd} &= \text{time realized by skater } i \text{ in season } t \text{ during tournament } k\\
&\quad\ \text{at distance } d;\\
rT_{itkd} &= \text{time } T_{iktd}\text{, reduced to a 500m time, realized by skater } i \text{ in year } t\\
&\quad\ \text{during tournament } k \text{ at distance } d\text{, i.e.,}\\
rT_{itkd} &= \frac{|d|}{500} T_{itkd}\text{, with } |d| = \text{ the length of distance } d \text{ in meters.}
\end{aligned}
$$

In the following sections, $T_{itkd}$ is called the *time* of skater $i$, and $rT_{itkd}$ is called the *r-time* of skater $i$. So r-times are the ’500m-times’ as used in the samalog system; see Section 2.2.3.

## 2.3 The development of skating times and participation numbers

Like most other sports, speed skating has evolved over the years. The current skaters are faster than their former colleagues. In Figure 2.2, for both men and women, the progress of the speed skating times is shown. The years referring to the seasons are depicted on the horizontal axis and the r-times on the vertical axis. A graph corresponds to a distance, and a point on a graph corresponds to a season and the median of all r-times realized in that season.

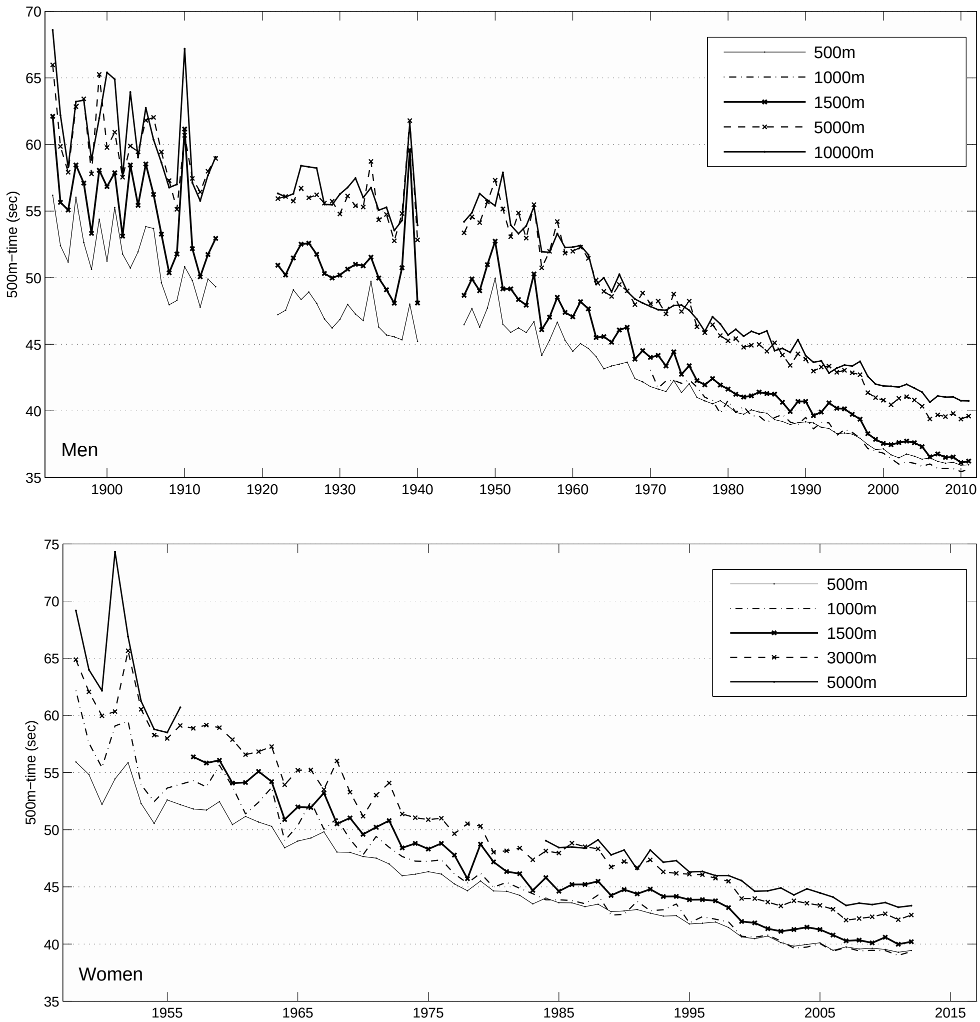



**Figure 2.2.** Progress in speed skating times, men and women

Figure 2.2 shows that the median of the r-times for all distances have decreased rather rapidly. For instance, in 1896 the median of all r-times of the 1500m is 58.13

seconds. About hundred years later, in 2006, this r-time has dropped to 36.38 seconds, an increase in speed of 38%. Similar results are seen for women. In 1956, the r-time of the 1500m for women was 56.2 seconds. Now, in 2006, they complete the same distance with an r-time of 40.5 seconds, an increase in speed of 27%.

Another characteristic of the graphs of Figure 2.2 is that also the fluctuations decrease over time. Until 1960, for both men and women, all graphs fluctuate strongly. The main reason of this phenomenon is the fact that before 1960 the number of competitors at the two allround tournaments was lower and less constant (see Section 2.3.5) than nowadays. The innovation in rinks also contributed to the decreasing fluctuations (see Section 2.3.2). In the next sections we will explain in more detail the main factors that caused the increase in speed, and the decrease in fluctuations.

### 2.3.1 General performance level

In all athletic sports, performances become better and better. The progression in performance can partly be explained by the extension of the human capacity, due to better training, nutrition and increasing participation (see, Robinson and Jonathan (1995)). However, in speed skating, a second factor plays a significant role, namely the technological innovations and improvements. In sports like running, the world records improve rather monotonically, Figure 1.1 illustrates this phenomenon, whereas in speed skating, new technological innovations have caused major jumps in the world record. In fact these innovations improve the performances of all skaters.

In all sports top athletes perform at a certain level and several developments will influence this level over time. We will use the term *General Performance Level* (GPL) of a sport at a certain time instance, and describes it for the time being as follow:

$GPL =$ The average level of performance of top athletes at a certain moment.

There are several ways to measure or indicate such performance level but the main goal in this research is to use it to compare different time eras and thereby look more into the changes and relative level than the absolute level.

For track and field sports and speed skating the actual times or scores of a fixed set of top athletes can be taken as GPL. However for teams sport or sports where athletes compete directly against each other it is harder to find a statistic. Match outcomes are relative statistics and as competing teams/players improve their defensive and offensive skills, it is more difficult to see an increase in the GPL. For example, Milanovic (2005) showed that in soccer the goals difference per match of the top eight teams during World Cups over the last fifty years has decreased, not because the players have become worse, but play in total has improved.

In speed skating the GPL is influenced and improved by the following developments:

- technological innovations, see Section 2.3.2;
- maturation of the sport, see Section 2.3.3;

- population and participation changes, see Section 2.3.5.

These developments and their influence on the GPL of speed skating are described in the next sections.

### 2.3.2 Technological innovations

We define technological innovations as innovations that change the materials and the rinks in speed skating and improve the performance of all skaters. They can be divided into two groups, namely better equipment such as tight fit suits and klap-skates, and better rinks such indoor-rinks and artificial ice tracks.

**Rink and track improvements**

The improvements of the rinks started with the building of high altitude rinks. Due to the thinner and dryer air, the air resistance on high altitudes is low, so that skaters are able to go faster. The Davos rink in Switzerland and the Medeo rink in the former Soviet Union are classical examples. Especially in the period 1950-1970, mainly Soviet skaters set new world records at the Medeo rink. Besides the altitude, the location of the Medeo rink between the mountains occasionally caused special 'fall' winds, which served as a push in the back of the skater almost the complete lap. The Soviets organized contests when such winds were expected.

After the Medeo rink was destroyed by a rock avalanche, it was rebuild in 1972 with 'artificial' ice, which is ice made by cooling machines. The first artificially refrigerated rink was established in 1958, namely the Nya Ullevi rink in Gothenburg. This new technique has resulted into a complete new generation of rinks, especially at locations where the temperature during the winter period is less frequently below zero. This new technique also brought opportunities to improve the ice quality. By changing the composition of the water, one could lower the ice resistance on the blades. Nowadays, rinks technicians still experiment and try to find the optimal ice composition for their track.

A major innovation was the construction of indoor rinks. The first one was built in 1986 in Heerenveen, the Netherlands. One year later, the first indoor Olympic Winter Games for speed skating were organized in Calgary. The indoor rinks created better circumstances, enabling skaters to go faster. They also expelled the influence of changing weather circumstances, so competitions became more fair in this way.

The introduction of artificial ice and indoor rinks not only improved the skating times, they are also responsible for the damping of the fluctuations of the graphs in Figure 2.2. The peaks and fluctuations in the early periods are mainly caused by the changing circumstances between the tournaments in each season. With artificial ice and indoor rinks more equal circumstances are created, and nowadays, the influence of the weather is almost completely eliminated. As a result the fastest times are only skated on high altitude indoor rinks with low air pressure.

**Innovations in equipment**

The improvement of the skates started at the end of the 19th century. The Norwegian skater, Axel Paulsen was the first skater who screw the iron blades onto his boots. Short after that, fellow-countryman, Harald Hagen introduced metal tubes in which the blades are clipped and this is still the basis of the modern skate. Around 1900, skaters, like the Dutchman Coen de Koning, already experimented with the thickness of their blades to improve their results.

The first important improvement in the skater's equipment was however the introduction in 1976 of the so-called tight fit suit by the Swiss speed skater Franz Krienbuhl. This suit fits very tight to the body and is more aerodynamic than the woolen jerseys. Nowadays, skating suits are tested in wind tunnels. In 2002, skaters started using strips attached to the caps and the arm sections of the suits. These strips should cause better aerodynamic air flows, although not all skaters are convinced of its working.

The innovation introduced in 1996 was a true revolution in the skating sport. The Dutch scientist Gerrit Jan van Ingen Schenau introduced a new type of skates, the so called klapskate. This new type of skate enabled practically all skaters to improve their personal records, and all world records were broken in a short period of time. Figure 2.2 clearly shows that since 1996 all skating times are improved significantly. In the period of five years after the introduction, it seems that the full effect of the klapskate was still not entirely exploited. Almost every year new world records were set and all skaters kept improving their performances. One of the reasons for this could be the necessary change in skating technique due to the working of the klapskate. As described in Houdijk *et al.* (2000), the push-off technique between the conventional skates and the klapskates is different. So the first users of the klapskate, who grew up with conventional skates, had to unlearn the traditional way of skating and to learn the new klapskate technique. The current generation of top skaters only practiced on klapskates, and therefore their technique is better adjusted. The training methods are now also adjusted to the use of the klapskate.

The effect of the klapskate on world records is discussed in Kuper and Sterken (2003). In this paper, the world records in 2006 are predicted based on the gains with the klapskate in the first years after the introduction. All predictions turned out to be too pessimistic. The effect of the klapskate seemed to be underestimated, mainly because the full potential of the klapskate was not yet completely used.

**Influence of innovations on world records**

The influence of rink and equipment improvements are best seen in the progress of the world records. Figure 2.3 shows, for both male and female skaters, the progress of the world records on all distances. The graphs of the world records show of course a decreasing trend and the large drops are mainly caused by technological innovations. For example, Figure 2.3 shows that in 1996, with the introduction of the klapskate, all world records dropped. The effects of rink improvements are less visible in the graphs, because these improvements are introduced not at one moment of time, but rink by rink over the years. During the last ten years almost all world records are skated in Calgary and Salt Lake City where the highest altitude indoor rinks are located.

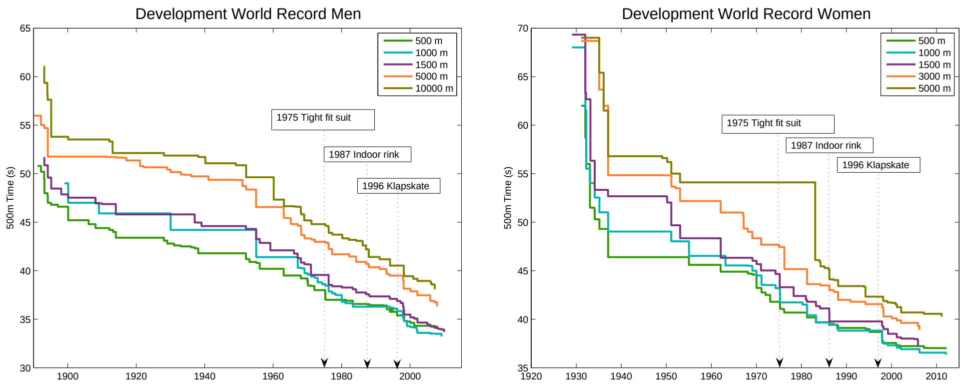



**Figure 2.3.** Progress of world records

## 2.3.3 Maturity of the sport

The rink and equipment innovations have caused ’jumps’ in the GPL of speed skating. After the introduction of these innovations, all skaters benefit from it and all are able to improve their performance. Figure 2.2 shows that there are not only ’jumps’, but that there is also a smoothly decrease in speed skating times. The following developments contribute to the continues improvement of the performances.

### Improving training methods

By improving existing training methods and increasing the knowledge about the effects of these methods, skaters are more efficient and effective in applying training techniques. Nowadays, talented skaters are earlier recognized, better accompanied and brought faster to the top. Only the most talented and best trained athletes take part in international tournaments, and with a larger group to select from, the average performance level of participating skaters has increased.

Introducing new training methods also has given a boost on the output of the skaters. A classical example concerns the American speed skater Eric Heiden, who combined his extraordinary talent with a revolutionary way of training (see Heiden and Testa (2008)). Heiden dominated the skating world for three years and gathered five gold medals during the 1980 Winter Olympics.

**Specialization**

Athletes that can focus on one event will perform better than if they have to divide their training assets to multiple events. In speed skating, the introduction of new tournaments, like the WCC and the WSDCh (see Section 2.2.2) have led to more specialization. Besides the allround tournaments, there are now tournaments in which skaters can win prices on one specific discipline.

**Financial resources**

Finally, the introduction of professionally sponsored skating teams has improved the performance. Already in 1973 a small group of skaters, including world champion and three times Olympic champion Ard Schenk, started their own professional skating league in order to generate more income. The experiment collapsed already in 1974. In 1995, Rintje Ritsma founded the first professional sponsor team with as sponsor Sanex, soon followed by other professional skating teams. The financial resources of the sponsor teams made it possible for skaters to train full time, while talented skaters are given the opportunity to start at early ages.

Developments such as specialization, better training methods, and team sponsoring have a continuous impact on the performances. In this chapter we will relate them to the maturity of the sport; the more mature a sport gets, the better the performances will be. The rate of growth of the GPL, caused by the development of the maturation process, will be called the *maturity level*.
In the following section we will discuss a general hypothesis made by Stephan Jay Gould regarding the maturity level and the influence on performance.

### 2.3.4 Gould's hypothesis

Gould discusses the disappearance of 0.400 baseball hitters, i.e., of baseball players that are able to hit an average of over 40% of the balls during one season (see Gould (1996)). Such hitting averages were common place in the twenties and thirties of the twentieth century, whereas these averages did not appear anymore since 1941. For a long time this phenomenon was explained by the fact that the performance level of the baseball players must have gone down. Gould suggested an opposite explanation, namely that the level has increased. The disappearance of 0.400 hitters in his view is caused by the fact that the performance of the pitchers and the field players has improved along with that of the batters. The hitting averages of all players stayed more or less constant, but because the average performances of both hitters and pitchers has increased over time, it has become harder and harder for top hitters to achieve such high hitting averages and to become so much better than their opponents. Gould summarized his findings in the following general hypothesis:

**Gould's Hypothesis.** *Complex systems improve when the best performers play by the same rules over extended periods of time. As systems improve they equilibrate and variation decreases.*

So, if we see sport as a complex system, the hypothesis indicates that as the sport matures, all top athletes perform better and the differences in performances between the top athletes decrease (more competition). Gould explains this by depicting the performances of the professional baseball population as a normal distribution in which the best performances lie at the right tail of this distribution. He claims that the performances of all sportsmen are bounded by certain limitations and will never surpass a fictitious boundary. So, as the sport matures, two things happen. First, the population of top athletes increases, and due to better knowledge and training methods, the performances of these top athletes increase as well, with the result that more and more athletes will come closer to the boundary. Secondly, athletes further away from the fictitious limit can improve easier than athletes who already are at the top. As a result, the difference between top and average athletes narrows down as the sport matures, leading, in general, to an increase in competition level.

Gould based his results on data from baseball, also supported by Schmidt and Berri (2003) and similar results are found in basketball (see, Chatterjee and Yilmaz (1999)), cycling (see, Wieting (2000)) and in golf (see, Berry *et al.* (1999)). For track and field sports Gould's hypothesis can also be observed. Data from these sports show that the gaps between world records and average performances of top athletes become smaller and smaller as time elapses. However, in order to quantify these phenomena, more research is needed; see Section 2.10

Another observation is that at the start of a 'new' sport, performances improve rather quick and that as the sport gets 'older' the rate and magnitude of improvements will tome down. A nice example of the influence of the maturity on the GPL is seen when performances of men and women are compared. In most sports, women started to participate at a professional level many years after the men did, and, as a result, improvements of the women's records are much larger than those of the men during the same time period. So it looks that women catch up with the men, but will the gap between men and women disappear? Many have studied this phenomenon, see Whipp and Ward (1992), Thibault *et al.* (2010), Kuper and Sterken (2009) and Tatem *et al.* (2004).

Tatem *et al.* (2004) even estimated that in the long run, women would sprint faster than men. The winning times were extrapolated with a constant rate and based on this they predicted that women would outrun men in the future. However, they did not take into account the maturity effect of Gould's hypothesis, saying that the improvements become smaller and smaller. In Thibault *et al.* (2010) is shown that for many sports, including speed skating, the gap between the two genders is actually stabilized.

**Gould in speed skating**

Gould's hypothesis can also be used to explain a part of the development in speed skating times. Regarding the first part of Gould's hypothesis, speed skating is practiced more or less under the same rules since 1892, namely skaters compete on 400m rinks and the fastest skater is the winner. However, major changes, like the technological innovations and the introduction of new tournaments, may interfere with the Gould's hypothesis.

The decreasing trend seen in Figure 2.2 supports the statement regarding the 'improve' part in Gould's hypothesis, covering effects like better training, specialization and sponsoring. Gould also claims that differences between performances become smaller. This means that the maturity factor not only increases the performance, it also increases the competition between skaters. Better training methods, for example, have resulted in the fact that young talented skaters can catch up faster and easier with the top. The result is that more skaters compete at the top. Specialism, partly as a result of the introduction of new tournaments (see Section 2.2.2), has created the opportunity for skaters to focus more on their best discipline, increasing their chances of success, and, as more and more skaters specialize, the competition at the single distance events will increase. The introduction of sponsor teams have practically the same effect as better training methods: they bring young talented skaters faster to the top.

So, as performances are bounded on by the fictitious boundary and top skaters become closer and closer to it, there remains less and less space to improve, leading to an increase in competition. Besides the innovations, one of the most important factors that influences but also may disturb Gould's hypothesis is the size of the active population and the related participation level at tournaments. In Section 2.3.5, we will describe the developments in the skating population. In Section 2.6.3, we will investigate Gould's hypothesis for speed skating in more detail, and explain how to deal with population and participation changes.

### 2.3.5 Population and participation changes

Nowadays many skaters are active in the international skating circuit. As we will see in this section, new tournaments, aging, and globalization of the sport have led to an increase of participants at the international tournaments. We will also see that, in the last twenty years, skaters participate in an increasing number of international races during the season. This section only uses the data from male skaters, but all conclusions hold for women as well.

**The number of participants at tournaments**

The participation level at the various tournaments is depicted in Figure 2.4. For each tournament, the total number of competitors per season is plotted. The participation levels of the professional championships (see, Section 2.2.2) in 1973 and 1974 are not included. Figure 2.4 shows that the number of participants at the six tournaments

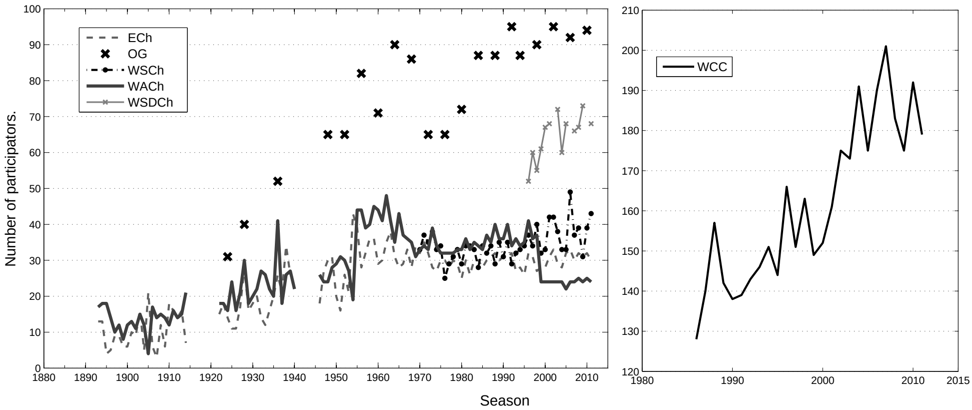



**Figure 2.4.** Total number of participants per tournament

not only varies between tournaments but changes over the seasons as well. The fluctuation over the years is partly caused by changing ISU regulations. For most tournaments, the ISU has established a time limit and a restriction on the number of participants. The rules itself have changed quite a bit over the years causing changes in the participation levels.

During the early years of speed skating, i.e., the period 1893-1914, the number of participants at WAChs fluctuates around fourteen, with the exception of 1889 and 1905 with eight and five participants, respectively. Due to the lack of good international transportation most participants came from the organizing country. This changed in the period 1922-1952, in which the number of participants increased to an average of 25. After that period, until 1999, this number never went below 32 with a maximum of 48 in 1962. In 1999, mainly due to the increased interest from television and the commercial world, the ISU restricted the number of participants to 24. The number of participants at the EChs has always been slightly below the number at the WAChs. Between 1893-1914 it varied between 4 to 16, and from 1950 until now a steady 30 skaters compete. Since 1948 the 10000m of both allround tournaments is only skated by a restricted number of participants, namely the best ranked skaters after three distances. From 1948 to 1955, only 12 skaters were allowed to skate the 10000m; in 1956 it was raised to 16, whereas in 1993 the ISU restricted it again to 12.

The total number of participants at the seven Olympic Winter Games of the last twenty years remained almost the same. Although in 1994 the restricted number of 10000m competitors dropped from 32 to 16, this is not explicitly seen in Figure 2.4 as most 10000m skaters also skated the 5000m. The number of skaters participating at WSDCh, which replaces the Olympics during non-Olympic years, is lower than the number of participants at Olympics. The ISU has set a maximum of 24 per distance, instead of the 32 for the Olympics. The number of participants at WSCh varies every year between 32 and 40, with the exception of 2006 when 49 skaters competed.

Finally, in the left panel of Figure 2.4, it can be seen that the number of skaters in

WCC increases from 128 in 1986 to 209 in 2007, merely due to the introduction of the so-called B-group in 1996, consisting of more or less sub top skaters.

### Number of observations per skater

In Section 2.3.5 we have analyzed the number of participants per tournament. However, to a large extend the WACh, the ECh, and later on the new tournaments have more or less the same participants. For example, in 2006, 11 of the 33 skaters were present at both the ECh and the WACh. Until the introduction of the new tournaments, for each distance, the number of international meetings between the skaters remained constant; after the introduction the figures changed somewhat.

In Figure 2.5 this is illustrated by the average number of international races per skater per distance, which is plotted for all skating seasons. Until 1970, the year of the first World Sprint Championships, for each distance, this average varies between one and two races with an average of 1.5 races per skater, because in the non-Olympic years of this period the ECh and WACh were the only international tournaments. So the 500m, 1500m, 5000m and 10000m could maximal be skated twice a year until 1970.

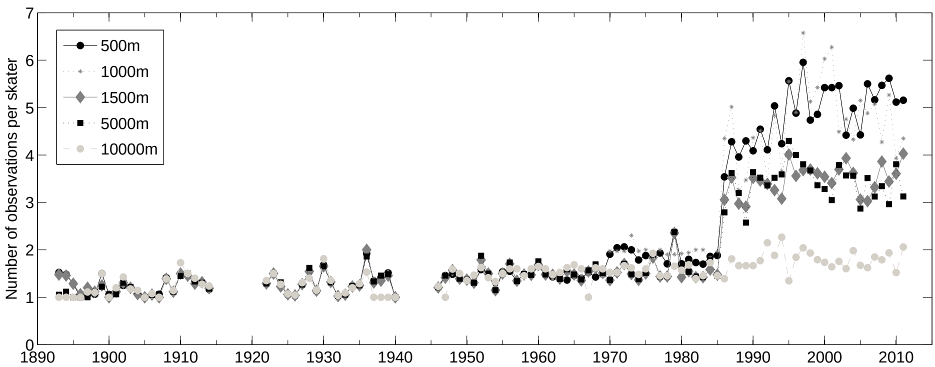



**Figure 2.5.** Observations per skater per distance

The introduction of the WSChs slowly divided the skate population into two groups, namely sprinters and allrouders, and slightly increased the average on the two shortest distances to two races per skater. This number increased further to an average of five, when the World Cup Competition was added to the skating calendar in 1986. Nowadays the sprinters skate nine to ten 500m's on average, whereas the pure allrounders remain to skate at most two 500m's. The WCC also increased the number of observations per skater on the 1500m and 5000m to an average value of 3.5. Note that the number of observations for the 10000m did not increase much after 1980, due to the fact that the 10000m's are skated only once or twice during a World Cup season.

In Table 2.2, the average number of observations per skater per tournament is given. Obviously, the allround tournaments have an average of around 3.5 observations, since each participant skates the 500m, 1500m, and the 5000m, and only the best 12 or 16 skate the 10000m, which is in most cases about 50% of the total group. For the WSCh this figure is somewhat higher, namely around 4, because, in principle, every sprinter skates both the 500m and the 1000m twice.

**Table 2.2.** Observations per male skater per tournament

| Tournament | # races | Tournament | # races |
|---|---|---|---|
| OG | 2.2 | WSCh | 3.8 |
| WACh | 3.4 | WSDCh | 2.0 |
| ECh | 3.4 | WCC | 7.5 |

Olympic participants skate on average 2.2 distances; most sprinters participate in the 500m and the 1000m, while long distance specialists participate in the 5000m and the 10000m. On the 1500m, skaters from both groups compete against each other. For the World Cups, we see from Table 2.2 that the average number of observations lies around 7 or 8 observations. However of all tournaments, the World Cup average fluctuates the most over the seasons and between the skaters. Especially between sprinters and long distance skaters the difference can be very large. During one World Cup weekend, sprinters skate, as in the WSChs, both the 500m and 1000m twice, whereas a long distance skater has only one distance race. The results show that, on average, sprinters skate around 10 through 12 World Cup races per season, and long distance skaters skate only 5 through 7 races.

**Countries and participation numbers**

In the previous two paragraphs, we have seen an increase in the number of active skaters and observations per skater over the years. Also the number of countries has increased. In Figure 2.6 two graphs are presented, namely, the total number of countries active in at least one of the six major tournaments per season (indicated by the black line), and the percentage of skaters from the six skating countries present at the WAChs (indicated by the gray dotted line). These 6 countries are Norway, Sweden, the Netherlands, USA, Finland, and the combination of Russia and the USSR, which are the top 6 best represented countries during WAChs over the years. The graph of the total number of active countries shows an upwards trend until 1999 when it starts to decline. In the period 1892-1914, around five to seven countries were active, whereas currently skaters from over 25 different countries compete. The decline after 1999 can be explained by the fact that the participation level at World Allround Championships was lowered from 32 to 24.

The overall increase in active countries also influenced the number of skaters per country present at WAChs. The gray dotted line of Figure 2.6 show that between 1892 and 1940 most skaters at the WAChs are from the top six counties. Especially Norway, Sweden and Finland are highly present. Due to the growing number of countries, the ISU allowed only five skaters per country in the period 1945-1972, and

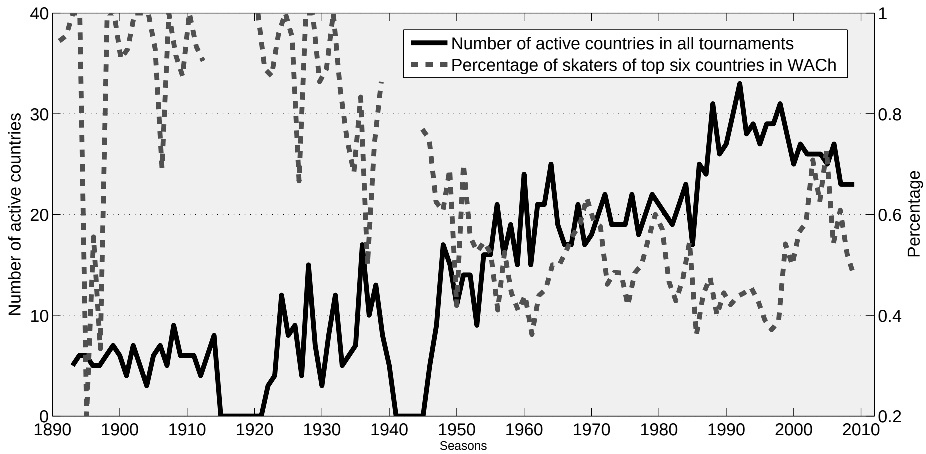



**Figure 2.6.** Number of countries active in speed skating (black line), and percentage of total participants at WACh of the top six countries (grey dotted line) for men.

after 1972 only four. Since 1972 no country is represented by more than five skaters.

Figure 2.6 also shows that in the last thirty years, the increase in numbers of countries (increasing trend of the black line) correlates negatively with the percentage of skaters from the top six countries (decreasing gray dotted line). As more and more countries become active, fewer places are available for the top countries. The number of active skaters per country is also growing steadily, around 1900 an active country had, on average, four active skaters in the international tournaments, where in 2010 this number has increased to seven active skaters. The introduction of the World Sprint Championships, the World Cup Competition and the World Single Distance Championships are the main reason for this increase.

The combination of the growing number of countries seen in Figure 2.6 and the growing average number of active skaters per country leads to a growth in the total number of active skaters. This may have, as Gould hypothesis describes, an impact on the GPL and competition level. Due to the fact that there are limited starting tickets for the major tournaments, only the very best compete and difference will become smaller. On the other hand, the limitation is also made per country. Some countries have many skaters that belong to the top, but can only sent their very best ones and let other countries sent less talented skaters to the tournament. So due to the country restrictions not always all the best skaters are present at the major tournaments. Meaning that the competition level is not as high as it could be and that the effect of the increasing in population with respect to Gould's hypothesis is harder to identify as we do not observe all top performances.

## 2.4 The ranking system

The objective of this chapter is to design a system that ranks skaters of all times. But how to compare skaters of all times? In order to answer this difficult question, we need to define a performance measure that allows for comparing performances of skaters from different time eras. We use a number of preliminary formulated criteria, used to 'judge' the final rankings.
This section will first list the disciplines for which we design the rankings and profile what defines the 'best' skaters. Secondly, we will formulate criteria for the performance measure, the ranking method and the ranking results, that need to be satisfied.

### 2.4.1 Speed skating disciplines

Most current world ranking systems rank skaters on their personal best times per distance. The Adelskalender, the oldest and most used 'all times' ranking system, uses the personal records of the four allround distances, which are aggregated into a ranking score (see Section 2.1.2). This ranking reflects the performances of the current skaters but is certainly not a ranking of all times.

We introduce a new system, namely the so-called *Universal Speed Skating* (USS) ranking. Both for men and women, the USS rankings include ranking lists for seven disciplines, namely for each distance one list, one sprint list containing the results of the 500m and 1000m, and an overall list containing all results.

The disciplines will be labeled by $L$, whereas $L$ is a set of distances $d \in D$ belonging to the discipline. For the male skaters, the set of disciplines is denoted and the defined as

$$\begin{aligned} DS_M &= SD_M \bigcup SP \bigcup OV_M \\ \text{with} & \\ SD_M &= \{500m, 1000m, 1500m, 5000m, 10000m\} \\ SP &= \{\{500m, 1000m\}\} \\ OV_M &= \{D_M\} \end{aligned}$$

and for women, the set of disciplines is denoted and defined as

$$\begin{aligned} DS_W &= SD_W \bigcup SP \bigcup OV_W \\ \text{with} & \\ SD_W &= \{500m, 1000m, 1500m, 3000m, 5000m\} \\ OV_W &= \{D_W\} \end{aligned}$$

So $L \in DS_M$ means that $L$ is either one distance $d$ ($L \in SD_M$) for the individual male lists or $L = 500m, 1000m$ for the sprint list or $L = D_M$ for the overall male list. In the above definition we dont explicitly distinguish between the inner and outer lane 500m and 1000m.

### 2.4.2 What defines a best skater?

A unique definition of 'best skater of all times' is hard to give. One could say that Sven Kramer, who has currently the lowest score on the Adelskalender, is the best skater, or maybe Eric Heiden, who won five gold medals on the Olympics, or even Oscar Mathisen and Clas Thunberg who both won five world titles as Kramer recently did.

One could also argue that the best skaters are the ones who dominated the skating sport for a longer period of time and were seen as 'unbeatable' in their time. Such 'best' skaters won most of the important tournaments, while these victories were achieved with great superiority. In Section 2.5, a summary of victories of these type of skaters is given, and in Section 2.5.3 existing ranking methods are compared. Subsequently, the shortfalls of these methods are discussed and it is shown that they do not satisfy our pre-conditions formulated in Section 2.4.3. In the USS-rankings, skaters will be judged on two aspects, namely important tournament victories and relative differences with opponents. So, not only the quantity of victories will be used, also the quality of these victories is taken into account.

### 2.4.3 Performance score

In order to rank the skaters, the USS-rankings will need a skater's performance score for the major international tournaments distance races. Absolute skating times (r-times) cannot be used since they are, as described in Section 2.3.1, influenced by technological innovations and the maturity of the sport, and therefore not comparable. However, absolute times do contain information regarding the (relative) quality of skaters which we want to use to rank the skaters. Therefore, the score used by the USS rankings, will be derived from the absolute r-times. It will measure the relative individual quality of the skater during a race and is called the *performance score*. For any distance, performance scores are 'comparable' entities over the years, where 'comparable' refers to the individual quality of the skater. Below a number of performance criteria are formulated, and based on previously discussed problems regarding the comparison of skating results, see Section 2.3 these are necessary conditions to make the performance scores comparable. There may be more criteria, however finding the 'best' set of criteria is not investigated and left for further research.

#### Performance criteria

The performance score of a skater is a transformed race time, that should satisfy the following criteria.

**PS 1** Performance scores are independent of technological innovations.

**PS 2** Performance scores are independent of the maturity level.

**PS 3** Performance scores are independent of the number of active skaters and competition levels.

**PS 4** Performance scores can differentiate between equal positions of different editions of the same tournament.

Criteria PS 1, 2 and 3 imply that r-times need to be corrected for both the introduction of technological innovations, the level of maturity of the sport, and the changing number of active skaters and the competition level. Technological innovations, such as the klap skate, make all skaters faster, not 'better'. The performance score of a skater should be defined and measured in such a way that the results of the 'old' skaters are comparable with the results of the current generation. We assume that, keeping in mind the competition level, winning in the early times could be as impressive as winning nowadays. This means that current world champions are not automatically 'better' than the old ones.

Criterion PS 2 is related to the hypothesis of Gould (see Section 2.3.4). According to Gould, skaters become faster and faster simply because 'the game is played'. In other words, the sport matures over time and so performances increase over time. In our ranking system we need to test for the influence of the maturity level on the performance score.

Criterion PS 3 refers to the facts described in Section 2.3.5 concerning the changing level of competition. In the early years of speed skating it was easier to win tournaments and stay at the top than nowadays. The performance scores should not be influenced by this fact.

We assume that if criteria PS 1, PS 2 and PS 3 are satisfied, we have obtained scores that are suitable to compare performance of skaters for any distance race. This fact can be formulated by the following pre-condition:

**Pre-condition 1.** *After finishing times are corrected for the influences of innovations, maturity level, and participation levels, then the performance scores of two skaters that have competed on different editions of the same tournament but on the same distance race are more or less equal if they finish on equal race positions.*

This pre-condition implies that, despite the fact that finishing times of skaters, who finished on equal positions on a certain distance race but in different years, are usually completely incomparable, their achievement and thereby quality should be seen as more or less equal. For example, the achievement of a skater finishing on the fifth place on the 500m of the Olympic Games of 1932 is considered to be in the same range as to the performance of the skater finishing at the fifth position on the 500m of the Olympic Games of 2006. However as we see the performance of each position in the same range, we have Criterion PS 4 which requires that we still can differentiate between these performances, for example, winners with a larger leads have a better performances score.

### 2.4.4 Ranking score criteria

Once for each skater the performance scores for all distance races are found, the scores of different tournaments should be aggregated to one ranking score and based on this ranking score the skaters are ranked. In the aggregation process of the performance scores, the resulting ranking score should satisfy the following criteria.

**RS 1** The impact of a performance score of a tournament on the ranking score depends on the importance of that tournament;

**RS 2** Performances scores of new tournaments should be included in the ranking score, but skaters from the period before its introduction should not have a (dis)advantage;

**RS 3** The ranking score of a skater should be independent of the length of that skater's career.

Criterion RS 1 indicates that tournaments have a certain hierarchical ordering regarding their importance: some tournaments have a higher status (to skaters, public, and media) than others. Winning a gold medal at the Olympic Winter Games is considered to be much more important than winning the same distance at a World Cup.

Criterion RS 2 refers to tournament changes as described in Section 2.2.2. Especially the new tournaments have led to more specialization. Nowadays, many skaters focus on these new tournaments, and if these tournaments are not included in the ranking, the current specialists have a disadvantage.

Finally, criterion RS 3 refers to the influence of the length of careers. The length of the career differs a lot between skaters. Where some are active only a few years in the international circuit, others are a whole decade. Skaters with long careers should not have a disadvantage for the fact that they did not belong to the absolute top at the beginning or the end of their career. A longer career should not automatically increase the ranking position as well.

### Ranking output criteria

Finally, we formulate a number of criteria regarding the final ranking lists. In Section 2.8.4 we will validate the criteria on the USS-rankings. The USS-rankings should satisfy the following output criteria:

**RO 1** Olympic and world champions are in the top of the USS-rankings.

**RO 2** Top ranked skaters are more or less uniformly distributed across the years.

**RO 3** Period colleagues are ranked correctly according to the tournament results in which they competed each other.

**RO 4** The relative ranking of retired skaters does not change when new skaters enter the rankings.

**RO 5** Experts opinions and other ranking methods should be consistent with the final ranking.

Criterion RO 1 refers to the fact that Olympic and world champions are present in the top of the lists. This criterion is quite obvious since champions belong to the best skaters of their time. It will also verify the pre-condition 2.4.3: all champions became number one during a tournament and as they are in the top of the lists, their performance scores must be more or less equal.

Criterion RO 2 refers to the fact that each period has its own champions and they should be present in the top of the ranking. Nevertheless, we may expect more skaters from last two decades in the sub top, because of the increased number of skaters and tournaments.

Any subset of an USS-ranking, related to a certain time period, should reflect the differences in average performance of the skaters of in that period. In other words, if skater A was 'better' than B in a certain time period, than A is ranked higher than B in the USS-ranking.

Each year new results and new skaters will enter and change the USS-rankings. Although retired skaters may be ranked lower, their relative positions will not change.

Finally, criterion RO 5 states that the USS-rankings should, to a certain extent, be consistent with the opinions of 'experts' and with the ranking from Section 2.5.

## 2.5 Best skater indicators

Nowadays, the speed of top skaters is higher than ever, but that does not directly mean that the current top skaters are 'better' than their former colleagues. Many 'old' champions are still remembered by their victories or impressive records. In this section we will list the skaters with the most tournament and championship victories (Section 2.5.1), and the ones with the most world records (Section 2.5.2).

### 2.5.1 Tournament winners

In Table 2.3 and Table 2.4, we have listed the most successful skaters on the four major tournaments. The tables contain the skater's names, the number of victories, and the seasons of the victories. In the heading of the table also the number of times the tournaments are organized so far, is given. For example, since 1893, the men's WACh are held 106 times. The first world champion allround was Jaap Eden, who won in total three world titles. Oscar Mathisen, Clas Thunberg and Sven Kramer all managed to win five times the WACh. Kramer won his last one in 2012, and in the near future we may expect that he becomes the first skater with six titles.

Most athletes use several Olympics to win more than three Olympic gold medals. For example, it took Thunberg two Olympics (1924, 1928) to win his five gold medals. Jevgeni Grisjin (1956, 1960, 1964) and Ivar Ballangrud (1928, 1932, 1936) even competed in three Olympic tournaments to win their four gold medals. Eric Heiden, however, collected his five medals during one Olympic tournament (1980) by winning all five distances.

Although the WAch and ECh are both held 106 times, the number of skaters who won the ECh three times or more is smaller than the ones with three or more world allround titles. This might be somewhat surprising because there is less competition during European championships, mainly due to the lack of North American and Asian skaters. Maybe top skaters are more eager to win WACh, which is always organized later on in the season.

The figures for women in Table 2.4 show that Gunda Niemann is the 'queen' of the allround tournaments: she won both allround tournaments eight times. Most sprint titles, six in total, are won by Karin Enke, who also won five allround world titles. Lidia Skoblikova has the most Olympic titles; she won six gold medals during the first two Olympics for women.

Table 2.3 and Table 2.4 present the champions with the most victories. According to the criteria from Section 2.4.4, one may expect these skaters in the top of the USS-rankings.

#### Winners per country

Table 2.5 lists the number of podium places per counties for all tournaments except for the World Cups, both for men and women.

The figures of Table 2.5 show that Norway and the Netherlands have dominated

**Table 2.3.** Most victories, men

**World Allround Championships, Men**
Held since: 1893, 106x

| Name | # | Years |
|---|---|---|
| Oscar Mathisen | 5 | 1908, 1909, 1912-1914 |
| Clas Thunberg | 5 | 1923, 1925, 1928-1930 |
| Sven Kramer | 5 | 2007-2010, 2012 |
| Rintje Ritsma | 4 | 1995, 1996, 1999, 2001 |
| Ivar Ballangrud | 4 | 1926, 1932, 1936, 1938 |
| Hjalmar Andersen | 3 | 1950, 1951, 1952 |
| Jaap Eden | 3 | 1893, 1895, 1896 |
| Eric Heiden | 3 | 1977, 1978, 1979 |
| Johann Olav Koss | 3 | 1990, 1991, 1994 |
| Ard Schenk | 3 | 1970, 1971, 1972 |
| Michael Staksrud | 3 | 1930, 1936, 1938 |

**European Championships, Men**
Held since: 1893, 106x

| Name | # | Years |
|---|---|---|
| Rintje Ritsma | 6 | 1994-1996, 1998-2000 |
| Sven Kramer | 5 | 2007-2010, 2012 |
| Ivar Ballangrud | 4 | 1929, 1930, 1933, 1936 |
| Clas Thunberg | 4 | 1922, 1928, 1931, 1932 |
| Hjalmar Andersen | 3 | 1950, 1951, 1952 |
| Rudolf Gundersen | 3 | 1901, 1904, 1906 |
| Oscar Mathisen | 3 | 1909, 1912, 1914 |
| Ard Schenk | 3 | 1966, 1970, 1972 |

**Olympic Winter Games, Men**
Held since: 1924, 21x

| Name | # | Years |
|---|---|---|
| Eric Heiden | 5 | 1980 |
| Clas Thunberg | 5 | 1924, 1928 |
| Ivar Ballangrud | 4 | 1928, 1932, 1936 |
| Yevgeni Grisjin | 4 | 1956, 1960, 1964 |
| Johann Olaf Koss | 4 | 1992, 1994 |
| Hjalmar Andersen | 3 | 1952 |
| Tomas Gustafson | 3 | 1984, 1988 |
| Ard Schenk | 3 | 1972 |

**World Sprint Championships, Men**
Held since: 1970, 43x

| Name | # | Years |
|---|---|---|
| Igor Zjelezovski | 6 | 1985, 1986, 1989, 1991-1993 |
| Eric Heiden | 4 | 1977-1980 |
| Jeremy Wotherspoon | 4 | 1999, 2000, 2002, 2003 |
| Kyou-Hyuk Lee | 4 | 2007, 2008, 2010, 2011 |
| Dan Jansen | 2 | 1988, 1994 |
| Sergej Klevtsjenja | 2 | 1996, 1997 |
| Akira Kuroiwa | 2 | 1983, 1987 |
| Valeri Muratov | 2 | 1970, 1973 |
| Erben Wennemars | 2 | 2004, 2005 |

# = the number of titles

**Table 2.4.** Most victories, women

**World Allround Championships, Women**
Held since: 1947, 66x

| Name | # | Years |
|---|---|---|
| Gunda Niemann | 8 | 1991-1993, 1995-1999 |
| Karin Enke | 5 | 1982, 1984, 1986-1988 |
| Atje Keulen-Deelstra | 4 | 1970, 1972-1974 |
| Inga Artamonov | 4 | 1957, 1958, 1962, 1965 |

**European Championships, Women**
Held since: 1947, 66x

| Name | # | Years |
|---|---|---|
| Gunda Niemann | 8 | 1989-1992, 1994-1996, 2001 |
| Anni Friesinger | 5 | 2000, 2002-2005 |
| Andrea Mitscherlich | 5 | 1983, 1985-1988 |
| Martina Sablikova | 4 | 2007, 2010-2012 |
| Atje Keulen-Deelstra | 3 | 1972-1974 |

**Olympic Winter Games, Women**
Held since: 1960, 14x

| Name | # | Years |
|---|---|---|
| Lidia Skoblikova | 6 | 1960, 1964 |
| Bonnie Blair | 5 | 1988-1994 |
| Claudia Pechstein | 4 | 1994, 1998 & 2002 |
| Karin Enke | 3 | 1980, 1984 |
| Yvonne van Gennip | 3 | 1988 |
| Gunda Niemann | 3 | 1992, 1998 |
| Marianne Timmer | 3 | 1998, 2006 |

**World Sprint Championships, Women**
Held since: 1970, 43x

| Name | # | Years |
|---|---|---|
| Karin Enke | 6 | 1980, 1981, 1983-1987 |
| Monique Garbrecht | 5 | 1991, 1999-2001, 2003 |
| Bonnie Blair | 3 | 1989, 1994, 1995 |
| Sheila Young | 3 | 1973, 1975, 1976 |

# = the number of titles

the allround men's tournaments. In the period 1893-1952, Norway was superior. In that period, with 45 tournaments, 25 world titles and 24 European titles went to a Norwegian skater. During the last twenty years, the Netherlands has taken over the supremacy. From 1983 on, the European allround title went only six times to another country than the Netherlands, and fifteen of the last twenty World Allround Championships titles went to a Dutch skater. The Dutch superiority in the World Allround Championships was partly interrupted by the victories of the Americans in 2004-2006. The dominance of the Netherlands in the last decade is also visible in the figures of the World Single Distances Championships. This tournament was organized for the first time in 1996, and the Dutch male skaters won a total of 89 medals, including 36 golden, whereas the nearest competing country, the USA, has

**Table 2.5.** Victories per country

| **Olympic Winter Games** | | | | | **World Allround Champ.** | | | | | **World Sprint Champ.** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Men | | | | | Men | | | | | Men | | | | |
| Country | Gold | Silver | Bronze | Total | Name | Gold | Silver | Bronze | Total | Name | Gold | Silver | Bronze | Total |
| NOR | 24 | 28 | 26 | 78 | NOR | 36 | 30 | 30 | 96 | USA | 8 | 6 | 8 | 22 |
| USA | 20 | 13 | 7 | 40 | NED | 31 | 15 | 25 | 71 | CAN | 6 | 9 | 3 | 18 |
| NED | 15 | 21 | 20 | 56 | FIN | 9 | 10 | 3 | 22 | RUS | 6 | 2 | 0 | 8 |
| URS | 12 | 10 | 9 | 31 | URS | 8 | 13 | 10 | 31 | KOR | 6 | 3 | 2 | 11 |
| SWE | 7 | 4 | 5 | 16 | USA | 8 | 5 | 5 | 15 | SOV | 4 | 6 | 5 | 15 |
| FIN | 6 | 6 | 7 | 19 | SWE | 3 | 3 | 6 | 12 | NED | 4 | 3 | 9 | 16 |
| Women | | | | | Women | | | | | Women | | | | |
| URS | 12 | 7 | 10 | 29 | URS | 24 | 25 | 20 | 69 | USA | 10 | 12 | 7 | 29 |
| NED | 12 | 8 | 6 | 26 | GER | 12 | 12 | 5 | 29 | GDR | 10 | 6 | 3 | 19 |
| USA | 11 | 10 | 12 | 33 | GDR | 10 | 8 | 5 | 23 | GER | 8 | 3 | 5 | 16 |
| GER | 11 | 13 | 8 | 32 | NED | 10 | 7 | 15 | 32 | URS | 3 | 4 | 3 | 10 |
| GDR | 6 | 11 | 8 | 25 | NOR | 4 | 5 | 8 | 17 | CHN | 2 | 3 | 1 | 6 |
| CAN | 5 | 8 | 5 | 18 | FIN | 3 | 4 | 2 | 9 | CAN | 2 | 2 | 4 | 8 |

| **European Champ.** | | | | | **World Single Dist. Champ.** | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Men | | | | | Men | | | | |
| NOR | 38 | 37 | 33 | 108 | NED | 36 | 29 | 24 | 89 |
| NED | 29 | 24 | 25 | 78 | USA | 10 | 4 | 9 | 23 |
| SOV | 10 | 7 | 8 | 25 | NOR | 7 | 9 | 5 | 21 |
| SWE | 10 | 3 | 8 | 21 | JAP | 7 | 6 | 5 | 18 |
| FIN | 7 | 9 | 6 | 22 | CAN | 6 | 9 | 10 | 25 |
| RUS | 5 | 1 | 5 | 11 | ZKO | 3 | 3 | 2 | 8 |
| Women | | | | | Women | | | | |
| GER | 14 | 13 | 4 | 31 | GER | 36 | 29 | 14 | 79 |
| GDR | 8 | 6 | 5 | 19 | CAN | 12 | 9 | 17 | 380 |
| NED | 6 | 12 | 15 | 33 | NED | 11 | 17 | 15 | 43 |
| URS | 4 | 4 | 8 | 16 | TSJ | 5 | 3 | 0 | 7 |
| TSJ | 4 | 0 | 2 | 6 | CHN | 3 | 5 | 2 | 10 |
| AUT | 1 | 1 | 1 | 3 | USA | 1 | 2 | 5 | 8 |

to be satisfied with only twenty-three medals. In case of the World Sprint Championships the Americans and the Canadians have the lead. However, if the results of the former Soviet Union and Russia are combined, than this combination has the best sprinters.

The figures for women show a different pattern. Here the German skaters, and to a lesser extent the former Soviet skaters dominate the scene. The figures for the German women are even better if we combine the results of the former German Democratic Republic and West-Germany. The USA women have acted strongly at the World Sprint Championships and the Olympic Winter Games.

### 2.5.2 World records

Since the foundation of the ISU all world records, for both men and women are registered. Through March 2012, the men skated 285 world records on the five classical distances. The 500m world record is broken most often, namely 74 times. On the 1000m, 1500m, 5000m, and the 10000m the world record is broken 48, 56, 58, and 49 times, respectively. On a total of 2236 male skaters, who participated in an international tournament, 107 skaters are responsible for these 285 world records.
For the women these figures are somewhat lower. In the period 1930-2011, they set 190 world records. The 5000m world record is broken only 27 times because this distance was not part of the allround tournaments for a period of 27 years. The 5000m world record of Rimma Zhukova, skated January 23, 1953, held for almost thirty years until it was broken by Andrea Mitscherlich in 1983. The world record on the 500m for women is broken 43 times. For the 1000m, the 1500m, and the 3000m these figures are 44, 36, and 40, respectively. A total of 62 out of the 1037 international female skaters broke a world record.

In the Tables 2.6 and 2.7, the ten skaters with the most world records are listed. We recognize the typical allround skaters, such as Oscar Mathisen and Ard Schenk, who have set world records on both short and long distances, and the specialists like Gianni Romme and Jeremy Wotherspoon who only established world records on their specific disciplines.

**Table 2.6.** Top world record skaters, men

| Name | Period | Total | 500m | 1000m | 1500m | 5000m | 10000m |
|---|---|---|---|---|---|---|---|
| Oscar Mathisen | 1908-1916 | 14 | 4 | 1 | 4 | 2 | 3 |
| Ard Schenk | 1966-1972 | 10 | 0 | 3 | 3 | 2 | 2 |
| Peder Ostlund | 1893-1900 | 10 | 2 | 2 | 5 | 0 | 1 |
| Jeremy Wotherspoon | 1995-2003 | 9 | 2 | 7 | 0 | 0 | 0 |
| Fred Anton Maier | 1965-1968 | 8 | 0 | 0 | 0 | 4 | 4 |
| Johann Olav Koss | 1990-1994 | 8 | 0 | 0 | 1 | 5 | 2 |
| Gianni Romme | 1997-2000 | 7 | 0 | 0 | 0 | 4 | 3 |
| Yevgeni Grishin | 1955-1963 | 7 | 4 | 1 | 2 | 0 | 0 |
| Yevgeni Kulikov | 1975-1981 | 7 | 5 | 2 | 0 | 0 | 0 |
| Shani Davis | 2006 - 2011 | 7 | 0 | 2 | 5 | 0 | 0 |

**Table 2.7.** Top world record skaters, women

| Name | Period | Total | 500m | 1000m | 1500m | 3000m | 5000m |
|---|---|---|---|---|---|---|---|
| Catriona Le May | 1997-2001 | 10 | 8 | 1 | 1 | 0 | 0 |
| Gunda Niemann | 1990-2001 | 10 | 0 | 0 | 0 | 5 | 5 |
| Zofia Nehringowa | 1929-1935 | 9 | 1 | 2 | 2 | 2 | 2 |
| Tatyana Averina | 1974-1975 | 8 | 2 | 4 | 2 | 0 | 0 |
| Cindy Klassen | 2001-2006 | 7 | 0 | 2 | 3 | 2 | 0 |
| Stien Kaiser | 1967-1971 | 7 | 0 | 2 | 1 | 4 | 0 |
| Christa Rothenburger | 1981-1988 | 6 | 4 | 2 | 0 | 0 | 0 |
| Karin Enke | 1982-1987 | 6 | 1 | 2 | 2 | 1 | 0 |
| Laila Schou Nilsen | 1935-1937 | 6 | 2 | 1 | 1 | 1 | 1 |
| Natalya Petrusyova | 1976-1983 | 6 | 0 | 3 | 3 | 0 | 0 |

In Figure 2.7, the periods in which the world records are skated are depicted. For each period of five years the number of world records per distance are shown. The figure shows the high number of world records in the first five years (1890-1895 for men, and 1930-1935 for women). After the introduction of the 'artificial' ice in 1958, a significant increase in the number of world records can be observed. The high peaks in the period 1970-1975 are due to the many attempts of the Soviet male skaters at Medeo, especially at the short distances (see also Section 2.3.2). In this period the 500m world record was broken fourteen times, and the 1000m world record six times.

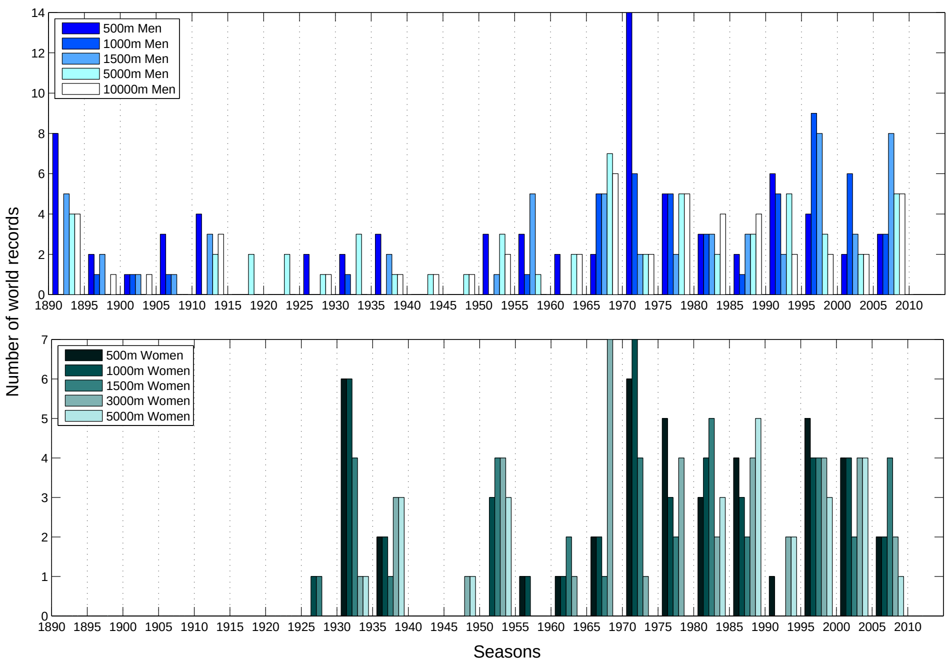



**Figure 2.7.** Number of world records per period

The introduction of the klapskates in 1996 is partly visible in Figure 2.7 as we see an increase in world records. In the period 1995-2000, the number of improvements is equally distributed over the distances for women. For the men, the world records of the 1000m and 1500m are broken more frequently in the period 1995-2006. The world records of the 1000m and the 1500m are broken nine and eight times respectively, the world records on the 500m, the 5000m and the 10000m only three, three, and two times respectively.

In the period 2005-2010, the world records on the longest distance are under fire. The world record on the 10000m was broken five times, whereas Sven Kramer is responsible for the last three improvements. In the previous two periods of five years the world record on the 10000m was only broken twice.

The number of world records skated in a period is heavily influenced by the introduction of new technology. Skaters who are active in that period benefit from the introduction and therefore world records are not useful for the comparison of performances.

### 2.5.3 Rankings based on 'medals'

In this section we discuss a number of specific ranking methods. These ranking methods are based on winning medals. So we can only rank skaters who finished within the best three on one of the distances or tournaments. The main problem in these type of rankings is how to compare and add up the types of medals obtained during different events. For example, how to compare Olympic titles with world titles?, and how to compare first places with second or third places?

We present two types of ranking methods, namely a *weight ranking* and a *domination ranking*. The goal of both ranking methods is to aggregate incomparable *performance dimensions* (in our case the various medals) into one ranking score. One performance dimension represent the number of times a skater became first, second or third in a tournament ranking or distance race ranking. For example, the number of times a skater finished first on the 500m of the WACh. Within a single performance dimension skaters can be compared directly: one can just order the skaters from high to low. However, in general we have more than one performance dimension and to aggregate more performance dimensions certain rules need to be introduced. Before we formulate the rules of the ranking models we introduce the performance dimensions used in the ranking models.

We will use nine dimensions, namely the number of gold, silver, and bronze medals on Olympic distance races[2], the number of podium places on World Allround Championships, and the number of first, second, and third places on individual distances during the World Allround Championships (WAChD). Other tournaments cannot be used since, either they are not accessible for all skaters (European Championships), or they are not organized the complete period (World Cups, World Sprint Championships, and World Single Distances Championships).

In Table 2.8, we have listed the first twenty male skaters with the highest number of total observations on the nine performance dimensions. The table shows that Eric Heiden won five Olympic titles, three WACh titles, and ten distance races during the WACh. Furthermore, he was one time the number two of the WACh final ranking, one time the number two and two times the number three on a distance race of the WACh. In the following section, these performance dimensions will be used to rank skaters.

#### Rankings based on weights from the first three positions

The most common used ranking system in sports is the so-called weight ranking. In this system, weights are assigned to each performance dimension, and these di-

[2]The Olympic medals rewarded for the allround results in 1932 are not included.

**Table 2.8.** Number of medals won

| | | Performance dimension | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Championship | OG | | | WACh | | | WAChd | | | Total |
| | Position | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 | |
| | Skater | | | | | | | | | | |
| 1 | Ivar Ballangrud | 4 | 2 | 1 | 4 | 4 | 3 | 15 | 10 | 6 | 49 |
| 2 | Clas Thunberg | 4 | 1 | 1 | 5 | 2 | 0 | 13 | 6 | 3 | 35 |
| 3 | Rintje Ritsma | 0 | 2 | 3 | 4 | 2 | 3 | 3 | 8 | 6 | 31 |
| 4 | Ard Schenk | 3 | 1 | 0 | 3 | 2 | 2 | 8 | 3 | 8 | 30 |
| 5 | Johann Olav Koss | 4 | 1 | 0 | 3 | 1 | 1 | 7 | 5 | 4 | 26 |
| 6 | Ids Postma | 1 | 1 | 0 | 2 | 4 | 1 | 7 | 5 | 5 | 26 |
| 7 | Knut Johannesen | 2 | 2 | 1 | 2 | 1 | 0 | 7 | 5 | 3 | 23 |
| 8 | Kees Verkerk | 1 | 3 | 0 | 2 | 0 | 3 | 7 | 6 | 1 | 23 |
| 9 | Sten Stensen | 1 | 1 | 2 | 1 | 2 | 1 | 6 | 5 | 4 | 23 |
| 10 | Eric Heiden | 5 | 0 | 0 | 3 | 1 | 0 | 10 | 1 | 2 | 22 |
| 11 | Sven Kramer | 1 | 1 | 0 | 4 | 0 | 2 | 10 | 1 | 2 | 21 |
| 12 | Bernt Evensen | 1 | 1 | 1 | 2 | 1 | 3 | 3 | 7 | 2 | 21 |
| 13 | Yevgeni Grishin | 4 | 1 | 0 | 0 | 0 | 2 | 6 | 3 | 2 | 18 |
| 14 | Shani Davis | 2 | 2 | 0 | 2 | 1 | 1 | 6 | 2 | 2 | 18 |
| 15 | Fred Anton Maier | 1 | 2 | 1 | 1 | 0 | 1 | 2 | 5 | 2 | 15 |
| 16 | Hjalmar Andersen | 3 | 0 | 0 | 3 | 0 | 0 | 6 | 2 | 0 | 14 |
| 17 | Gianni Romme | 2 | 1 | 0 | 2 | 0 | 0 | 4 | 1 | 1 | 11 |
| 18 | Gaetan Boucher | 2 | 1 | 1 | 0 | 0 | 0 | 2 | 1 | 2 | 9 |
| 19 | Tomas Gustafson | 3 | 1 | 0 | 1 | 0 | 0 | 1 | 2 | 0 | 8 |
| 20 | Jochem Uytdehaage | 2 | 1 | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 6 |

OG = Olympic Games, WACh = World Allround Championship, and WAChd = Individual distances during WACh.

mensions are aggregated by taking a weighted sum over all performance dimension values. If we have $K$ performance dimensions, then for each performance dimension $k = 1, \ldots, K$, and each skater $i$, we define

$$c_{ik} = \text{the score of skater } i \text{ on performance dimension } k,$$

and

$$w_k = \text{the weight of performance dimension } k.$$

The weighted ranking score of skater $i$ is denoted and defined as

$$WRS_i = \sum_{k=1}^{K} w_k c_{ik}.$$

The weights of the most important performance dimensions should have the highest values, because they should have the highest influence on the ranking score. Each medal type (see Table 2.8) has two dimensions, namely its 'color' (gold, silver, bronze), and its tournament (OG, WACh, and WAChD).

In Table 2.9 we list a couple of options for the weight values of the nine medal types of Table 2.8. Option 1 uses a linear weight schedule for both the podium places and the tournaments. The podium places are weighted by the ratio 3:2:1 for the first, second, and third position, respectively, and the tournaments are weighted by the ratio 10:5:1 for OG, WACh, and WAChD, respectively. The final weights for each dimension are found by multiplying both weights.

**Table 2.9.** Weights for each performance dimension

| | OG | | | WACh | | | WAChD | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 |
| Option 1 | 30 | 20 | 10 | 15 | 10 | 5 | 3 | 2 | 1 |
| Option 2 | 60 | 40 | 20 | 15 | 10 | 5 | 3 | 2 | 1 |
| Option 3 | 10000 | 10000 | 10000 | 100 | 100 | 100 | 1 | 1 | 1 |
| Option 4 | $10^{11}$ | $10^{10}$ | $10^{9}$ | $10^{8}$ | $10^{7}$ | $10^{6}$ | $10^{4}$ | $10^{2}$ | 1 |

In Option 2, the weights for the Olympic medals are doubled compared to Option 1. The Olympics are now four times as important as the WACh, mainly because the Olympics are organized only once in four years.

Option 3 makes no distinction between podium places (all three medals have the same weight). In practice this means that the total number of dimensions is reduced from nine to three, namely the number of Olympic medals, world championship podium places, and world championship distance podium places, and only the total number of medals counts.

The fourth option ranks the skaters in a hierarchical (or lexicographical) way. This ranking method is often used to rank countries based on the number of medals won during the Olympics. In the hierarchical ranking, objects are first ordered on the most important dimension. In case a set of objects has the same score, the order within this set is determined by the next important dimension, and this process will continue either until all objects are ordered, or if all dimensions are considered. Our weight ranking score will order the skaters in a hierarchical way if the weights satisfy the following condition:

$$w_k \geq \prod_{j=k+1}^{K} \left\lceil \max_i c_{ij} \right\rceil ,$$

where $k = 1$ is the most important dimension and $k = K$ the least important dimension. The notation $\lceil x \rceil$ means the smallest integer not less than $x$. In Option 4, the hierarchical ranking weights for all dimensions are rounded up to the nearest tenfold such that the total score directly reflects the score on each dimension (compare the figures of Table 2.8 with the scores in Option 4 of Table 2.10).

**Results** Table 2.10 presents the ranking scores of the first twenty skaters for the four options of Table 2.9. The ranking scores of Options 3 and 4 are split in such a way that one recognizes the number of medals won on each tournament or distance race. For example, the score 71131 of Ballangrud in Option 3 means that he has seven Olympic medals, eleven WACh podium places, and thirty-one WACh distance race podium places. In Option 4 the numbers are split into first, second, and third places (the '-' separates the tournaments). For example, the score 220-211-06-02-02 of Davis means that he won two Olympic gold medals, two Olympic silver medals, no Olympic bronze medals, two times the WACh, one time a WACh second place, one time WACh third place, six WACh distance race first places, two WACh distance race second places, and two WACh distance race third places.

In three of the four rankings, Ballangrud and Thunberg occupy the first two places; only in Option 4 they are beaten by Heiden, who is the only skater with five gold Olympic medals. Ballangrud scores higher than Heiden on all other eight dimensions. Heiden's career was pretty short, namely five seasons, in comparison to the sixteen of Ballingrud. Most other top ranked skaters have longer careers and so more opportunities to win medals. Clearly the length of a skater's career has influence on the ranking positions.

Also the competition level may have influenced the rankings. Heiden won five gold medals during one Olympic Games which is nowadays, due to specialization, an almost impossible achievement. Current skaters need two or more Olympics to collect five or more gold medals in order to beat Heiden. Lesser competition is probably also an explanation for the relative high scores of Ballangrud and Thunberg: in the years 1920-1940 it was easier to dominate for a longer period of time.

Finally, we observe that Sven Kramer and Rintje Ritsma, with respectively five and four WACh titles, do not appear in the rankings of Option 3 and 4. The reason is that they cannot compensate the small number of gold Olympic medals with their allround titles.

**Table 2.10.** Results weight models

| | **Option 1** Skater | Ranking score | **Option 3** Skater | Ranking score | |
|---|---|---|---|---|---|
| 1 | Ivar Ballangrud | 356 | Ivar Ballangrud | 71131 | (7-11-31) |
| 2 | Clas Thunberg | 299 | Clas Thunberg | 60722 | (6-07-22) |
| 3 | Eric Heiden | 239 | Rintje Ritsma | 50917 | (5-09-17) |
| 4 | Johann Olav Koss | 235 | Johann Olav Koss | 50516 | (5-05-16) |
| 5 | Ard Schenk | 223 | Eric Heiden | 50413 | (5-04-13) |
| 6 | Rintje Ritsma | 196 | Knut Johannesen | 50315 | (5-03-15) |
| 7 | Knut Johannesen | 184 | Yevgeni Grishin | 50211 | (5-02-11) |
| 8 | Yevgeni Grishin | 176 | Ard Schenk | 40719 | (4-07-19) |
| 9 | Sven Kramer | 175 | Bernt Evensen | 40612 | (4-06-12) |
| 10 | Kees Verkerk | 169 | Kees Verkerk | 40514 | (4-05-14) |
| 11 | Ids Postma | 161 | Sten Stensen | 40415 | (4-04-15) |
| 12 | Bernt Evensen | 160 | Roald Larsen | 40413 | (4-04-13) |
| 13 | Hjalmar Andersen | 157 | Fred Anton Maier | 40209 | (4-02-09) |
| 14 | Oscar Mathisen | 148 | Leo Visser | 40207 | (4-02-07) |
| 15 | Sten Stensen | 142 | Tomas Gustafson | 40103 | (4-01-03) |
| 16 | Tomas Gustafson | 132 | Gaetan Boucher | 40005 | (4-00-05) |
| 17 | Gianni Romme | 125 | Birger Wasenius | 30411 | (3-04-11) |
| 18 | Fred Anton Maier | 118 | Bart Veldkamp | 30317 | (3-03-17) |
| 19 | Shani Davis | 116 | Hjalmar Andersen | 30308 | (3-03-08) |
| 20 | Birger Wasenius | 114 | Chad Hedrick | 30209 | (3-02-09) |
| | **Option 2** Skater | Ranking score | **Option 4** Skater | Ranking score | |
| 1 | Ivar Ballangrud | 526 | Eric Heiden | 500310100102 | (500-310-10-01-02) |
| 2 | Clas Thunberg | 449 | Ivar Ballangrud | 421443151006 | (421-443-15-10-06) |
| 3 | Eric Heiden | 389 | Clas Thunberg | 411520130603 | (411-520-13-06-03) |
| 4 | Johann Olav Koss | 375 | Johann Olav Koss | 410311070504 | (410-311-07-05-04) |
| 5 | Ard Schenk | 333 | Yevgeni Grishin | 410002060302 | (410-002-06-03-02) |
| 6 | Yevgeni Grishin | 316 | Ard Schenk | 310322080308 | (310-322-08-03-08) |
| 7 | Knut Johannesen | 294 | Tomas Gustafson | 310100010200 | (310-100-01-02-00) |
| 8 | Shani Davis | 269 | Hjalmar Andersen | 300300060200 | (300-300-06-02-00) |
| 9 | Rintje Ritsma | 266 | Knut Johannesen | 221210070503 | (221-210-07-05-03) |
| 10 | Kees Verkerk | 259 | Shani Davis | 220211060202 | (220-211-06-02-02) |
| 11 | Hjalmar Andersen | 247 | Gaetan Boucher | 211000020102 | (211-000-02-01-02) |
| 12 | Tomas Gustafson | 242 | Gianni Romme | 210200040101 | (210-200-04-01-01) |
| 13 | Bernt Evensen | 240 | Jochem Uytdehaage | 210100020000 | (210-100-02-00-00) |
| 14 | Sven Kramer | 225 | Uwe-Jens Mey | 210000000000 | (210-000-00-00-00) |
| 15 | Sten Stensen | 212 | Erhard Keller | 200000000101 | (200-000-00-01-01) |
| 16 | Ids Postma | 211 | Kees Verkerk | 130203070601 | (130-203-07-06-01) |
| 17 | Gianni Romme | 205 | Bernt Evensen | 121213030702 | (121-213-03-07-02) |
| 18 | Fred Anton Maier | 198 | Fred Anton Maier | 121101020502 | (121-101-02-05-02) |
| 19 | Gaetan Boucher | 190 | Piet Kleine | 120101030202 | (120-101-03-02-02) |
| 20 | Jochem Uytdehaage | 181 | Julius Skutnabb | 120001000000 | (120-001-00-00-00) |

### Weight ranking for all positions

The weighted ranking method described in the previous section is based on the first three positions and is not able to rank all skaters. We now describe an extension of this system by including all positions, which we will call the all position weight ranking. In this ranking all positions of the OG's and WACh's are included and are labeled $OGp$, $WAChp$, $WAChDp$, where $p$ refers to the $p$-th position on that tournament or distance race. For example, dimension $OG22$ is the number of times the skater has reached position 22 during an Olympic Games. The weights for the dimensions are defined as follow.

$$w_{OGp} = 50\exp(\frac{1-p}{2}),$$

$$w_{WAChp} = 20\exp(\frac{1-p}{2}),$$

$$w_{WAChDp} = \exp(\frac{1-p}{2}).$$

The tournament ratio (OG;WACh;WAChD) is chosen to be 50:20:1, and each position within a tournament is worth $e^{-0.5} = 0.61$ times[3] the points of one position higher. So, position $p+1$ is worth 0.61 times the points of position $p$. Since this ratio is larger than 0.5, it follows that finishing twice on position $p+1$ is rewarded with more points than one time on position $p$, since $(2)(0.61) > 1$. Moreover, finishing one time on position $p$ and one time on position $p+2$ is better than twice position $p+1$, since $1 + (0.61)^2 > 0.61 \cdot 2$. The negative exponential formula is used for decreasing the weight point differences between consecutive positions. Most points are given to the highest positions and the difference between high positions is much larger than between lower positions. For example, the difference between $OG2$ and $OG3$ is 11.9 points, whereas the difference between $OG9$ and $OG10$ is only 0.4 points.

The ranking results based on these weights are given in Table 2.11. In column three through five, the scores without the tournament multipliers are given. The total score (column "Total") is calculated by multiplying the third column by 50, the fourth by 20, and the fifth by 1. Ballangrud, Thunberg, and Heiden are the top three. Compared to Option 1 from Table 2.10, almost no difference is observed within the first twenty positions. This is to be expected because of the negative exponential in the weight function: the first three positions have the highest weights and are not easily compensated by lower positions.

### Domination ranking score

The results of the weight ranking models depend on the choice of the weights for the performance dimensions. By choosing weights of the performance dimensions one creates an cardinal ranking for the performance dimensions, i.e., both an ordering and a relative difference between each dimension has to be specified. For example,

[3]The ratio between position $p$ and $p+1$ is equal to $\frac{\exp(\frac{1}{2}-\frac{p}{2})}{\exp(\frac{1}{2}-\frac{p}{2}-\frac{1}{2})} = \frac{\exp(\frac{1}{2}-\frac{p}{2})}{\exp(-\frac{p}{2})} = e^{\frac{1}{2}} = 0.61$.

**Table 2.11.** Results all positon weight ranking.

| | Skater | OG | WACh | WAChd | Total |
|---|---|---|---|---|---|
| 1 | Ivar Ballangrud | 5.0 | 7.9 | 25.6 | 432.1 |
| 2 | Clas Thunberg | 4.4 | 6.3 | 19.8 | 367.8 |
| 3 | Eric Heiden | 5.0 | 3.7 | 11.7 | 339.1 |
| 4 | Johann Olav Koss | 4.7 | 4.0 | 11.7 | 324.7 |
| 5 | Ard Schenk | 3.4 | 5.0 | 14.0 | 282.7 |
| 6 | Rintje Ritsma | 2.6 | 6.6 | 13.1 | 275.5 |
| 7 | Knut Johannesen | 3.6 | 3.3 | 11.8 | 258.8 |
| 8 | Shani Davis | 3.3 | 3.1 | 9.9 | 235.4 |
| 9 | Kees Verkerk | 3.1 | 3.4 | 12.0 | 234.8 |
| 10 | Yevgeni Grishin | 4.2 | 0.7 | 8.9 | 233.9 |
| 11 | Hjalmar Andersen | 3.1 | 3.2 | 8.6 | 226.6 |
| 12 | Tomas Gustafson | 3.7 | 1.4 | 3.9 | 214.8 |
| 13 | Ids Postma | 1.7 | 4.8 | 13.1 | 196.0 |
| 14 | Sten Stensen | 2.3 | 3.3 | 11.7 | 194.1 |
| 15 | Sven Kramer | 1.7 | 4.7 | 12.4 | 190.1 |
| 16 | Bernt Evensen | 2.0 | 3.7 | 10.4 | 183.4 |
| 17 | Bart Veldkamp | 2.4 | 2.1 | 15.1 | 177.5 |
| 18 | Roald Larsen | 2.0 | 3.2 | 12.0 | 175.7 |
| 19 | Gianni Romme | 2.6 | 2.0 | 5.2 | 175.5 |
| 20 | Chad Hedrick | 2.5 | 1.9 | 5.8 | 169.9 |

one needs to specify how much more important an OG is with respect to a WACh and how many times more a gold medal counts compared to a silver medal.

In Cherchye and Vermeulen (2006), a ranking method is presented that only requires an ordinal ordering of performance dimensions, called domination ranking, and therefore disregards weights. This study applies the ranking method to rank Tour de France cyclists from the period 1953-2004. If we apply this method, it means that the dimensions should only be ordered in such a way that dimension 1 is more important than dimension 2, dimension 2 more important than dimension 3, et cetera. We use the ranking method of Cherchye and Vermeulen (2006) for ranking skaters from the period 1924-2011. The performance dimensions used are the same nine dimensions as used in the weight ranking model; see Table 2.8. The ordinal order of the dimensions we choose to be as follows (where $\succ$ means 'more important than'):

$$OG1 \succ WACh1 \succ OG2 \succ WACh2 \succ OG3 \succ$$
$$WACh3 \succ WAChD1 \succ WAChD2 \succ WAChD3.$$

The domination ranking of Cherchye and Vermeulen (2006) is based on a so-called compensation principle, which is defined as follows:
For $k, n \geq 1$, let $k$ be the number of ordinal ordered dimensions and $n$ the number of objects to be ordered. Recall that $c_{ij}$ the score of object $i$ on dimension $j$. A function $V : \mathbb{R}^k \to \mathbb{R}$ is called an ordering domination function if it satisfies for each $i_1, i_2 \in$

$\{1, \ldots, n\}$ the following implication:

$$\sum_{j=1}^{l} c_{i_1 j} \geq \sum_{j=1}^{l} c_{i_2 j} \ \forall l \in \{1, \ldots, k\} \Rightarrow V(c_{i_1 1}, \ldots, c_{i_1 k}) \geq V(c_{i_2 1}, \ldots, c_{i_2 k}).$$

The above implication is called the *compensation principle*. The compensation principle states that one unit more of a higher dimension may compensate for one unit less of a lower dimension, but not vice versa. So, $i_1$ is ranked higher than $i_2$ if $i_1$ has a higher score on all dimensions, or can compensate lower dimensions with units from higher dimensions. In Cherchye and Vermeulen (2006), it is shown that the dominance metric defined by

$$I_{i_1,i_2} = \min_{l \in \{1,\ldots,k\}} \sum_{j=1}^{l} (c_{i_1 j} - c_{i_2 j})$$

satisfies the compensation principle and it follows that object $i_1$ dominates object $i_2$ if $I_{i_1,i_2} \geq 0$.

In our situation, we have that $k = 9$ (the nine performance dimensions), and $n = 375$ (the number of skaters with at least one score on at least one of the nine dimensions). Since Heiden is the only skater with five gold medals on the Olympic Games, he cannot be dominated by any other skater:; he has the highest score on the most important dimension OG1 (i.e., $l = 1$), which means that for each $i$ it holds that:

$$I_{i,Heiden} \leq (c_{i,1} - c_{Heiden,1}) \leq (c_{i,1} - 5) < 0$$

while, in order for skater $i$ to dominate Heiden, we should have had that $I_{i,Heiden} \geq 0$. Furthermore, Heiden cannot dominate Ballangrud, since Heiden's sum on the first three dimensions (5+3+0=8) is smaller than Ballangrud's sum (4+4+2=10), i.e. $I_{Heiden,Ballangrud} < 0$. Finally, we can see that Ballangrud is dominating Koss. Both skater's score on the first dimension is equal, but the sum of the scores of Ballangrud are always larger than the sum of scores of Koss, i.e. $I_{Ballengrud,Koss} = 0$.

The final ranking of skater $i$ is measured by the so-called *net-dominance metric*, denoted and defined as

$$NDM_i = \sum_{j=1}^{n} I^0_{i,j} - \sum_{j=1}^{n} I^0_{j,i},$$

i.e., $NDM_i$ is for each skater $i$ the difference between the number of skaters that are dominated by $i$ and the number of skaters that dominate $i$.

The $NDM_i$-values are calculated for the 375 skaters that have a positive score on one of the nine dimensions. In Table 2.12, the twenty skaters with the highest $NDM_i$-values are listed. The value of $NDM_i$ is given in the column 'Total score'. Note that $NDM_i \leq 375$. In the other columns, "1" indicates that the 'row' skater is dominated by the 'column' skater, while "0" means that this is not the case.

The number one, Ballangrud, dominates 372 skaters and is not dominated by any other skater. The two skaters that are not dominated by Ballangrud, are Thunberg, who has more gold Olympic medals plus world titles, and Heiden, who beats Ballangrud on the first dimension. Although, they are relative low in the ranking, it is remarkable that Rintje Ritsma and Oscar Mathisen are only dominated by Ballengrud. Ritsma and Mathisen have no Olympic gold medals, but have high scores on the second dimension, namely the World Allround Championships title, and therefore are hard to dominate. On the other hand, since Ritsma and Mathisen have no gold Olympic medal, they can never dominate an Olympic champion.

In Cherchye and Vermeulen (2006) it is claimed that the dominance ranking is more robust than the weight rankings. In our case the top 5 of the dominance ranking is equal to the top 5 of option 1 and 2 of the weight rankings. Furthermore, 18 of the 20 skaters in the top 20 of the all position weight ranking (Table 2.11) are also in the dominance ranking (Table 2.12).

**Table 2.12.** Results dominance ranking

| | Total score | Ivar Ballangrud | Clas Thunberg | Ard Schenk | Johann Olav Koss | Eric Heiden | Knut Johannesen | Kees Verkerk | Shani Davis | Yevgeni Grishin | Sten Stensen |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Ivar Ballangrud | 373 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Clas Thunberg | 370 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Ard Schenk | 364 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Johann Olav Koss | 364 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Eric Heiden | 362 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| Knut Johannesen | 355 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| Kees Verkerk | 347 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| Shani Davis | 346 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 |
| Yevgeni Grishin | 346 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 |
| Sten Stensen | 341 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| Hjalmar Andersen | 337 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| Bernt Evensen | 336 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 |
| Ids Postma | 333 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| Sven Kramer | 332 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| Fred Anton Maier | 328 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| Gianni Romme | 320 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 |
| Chad Hedrick | 317 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 |
| Piet Kleine | 316 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |
| Bart Veldkamp | 310 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| Rintje Ritsma | 310 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

1 (0) means that the 'row' skaters is (not) dominated by 'column' skater

### 2.5.4 Drawbacks of existing ranking methodologies

In Section 2.5.3, we have analyzed a number of options for the weight ranking system and the domination ranking system. Since these rankings are based on medals, they certainly select the best skaters of all times. Furthermore, we may observe that the methods satisfy both criteria PS 1 and PS 2 from Section 2.4.3. Results of single tournaments or distance races are hardly influenced by technological innovations and maturity factors and equal positions obtain the same score. Also by using different weights for the tournaments, Criterion RS 1 is also satisfied.

However, the methods have also some disadvantages. We will shortly point out that they do not satisfy the criteria PS 3, PS 4, RS 2, and RS 3.

First, observe that both the weight ranking and the domination ranking can only use tournaments that are organized during the entire observed period; using newly introduced tournaments is not fair in these rankings, because skaters from before an introduction did not have had the opportunity to participate. This means that criterion RS 2 is not satisfied.

Secondly, the weight methods do not distinguish between equal positions of tournament and distance races and there by fail to satisfy PS 4. Each victory obtains the same score, where it is likely that some victories are more impressive than others. A victory with a difference of one second on the runner-up is usually more impressive than a close victory with only 0.01 seconds ahead.

A third problem is the fact that the ranking methods do not take into account the participation and the competition, as required by PS 3. Is a victory in a tournament with eight participants equal to a victory with 32 participants. The same problem arises when we rank all skaters and each position obtains a weight. As tournament participations change, the assignment of weights to positions should not be constant. For example, a ninth place with, say, twenty-four participants is different from a ninth place with thirty-two participants. A final problem concerns the career lengths: some skaters have been active for only five years where others more than a decade. In longer careers more points can be gathered, resulting in higher ranking positions. This means that criterion RS 3 is not satisfied.

Before we introduce a ranking system without these drawbacks, which will be developed in such a way that it satisfies all PS and RS criteria, we give a short summary of a the the bridging model on which this model is based.

### 2.5.5 The Bridging Model

The problem of comparing the performances of athletes from different eras is thoroughly studied in Berry *et al.* (1999). In this paper, a 'statistical time machine' is constructed, in which is estimated how an athlete from one era would perform in another era. In Berry *et al.* (1999) is used the fact that careers of athletes overlap through the years. This phenomena is called bridging, and forms in this paper the bases for comparing athletes from different ears. Because there are bridges between athletes' careers, there is an advancing connection between the performances of athletes from the early years to the current athletes.

Berry *et al.* (1999) use additive models to estimate the following aspects: the individual innate qualities of the athlete, the effect of aging, and for each year/season the relative 'difficulty' of the sport. An important assumption in bridging models is that each sport has its own specific peak age, namely the age that each athlete has his peak performance. However, the 'height' of the peak and the way to and down the peak differ per athlete. The final rankings in bridging models are based on these estimates by determining for any athlete how he would perform at his peak age in a chosen benchmark year.

In Berry *et al.* (1999) the relative difficulties of the years are used as control variable. The control variables are used to correct for technological innovations and other changes that have influenced the performance of the athletes through the years. Furthermore, it is assumed that the individual quality of the athlete depends on the decade in which he is born. In the underlying distribution, used to model the individual ability, the mean and the variation depend on this decade. In this way, a correction is applied for Gould's maturity effect that influences both the performance itself and the variance of the performance.

The model in Berry *et al.* (1999) reads as follows:

$$P_{it} = q_i(d_i) + y_t + f_i(a_{it})$$

with $P$ a chosen seasonal performance score of athlete $i$ in year $t$, $q_i(d_i)$ the ability of athlete $i$, with $q_i(d_i) = N(\mu_i(d_i), \sigma^2(d_i))$ and $d_i$ the decade in which athlete i was born, $y_t$ the relative difficulty of year $t$, $f_i$ the aging effect of athlete $i$, and $a_{it}$ the age of $i$ in $t$. They use Markov chain Monte Carlo algorithms to calculate the posterior distributions and estimate all the parameters.

Table 2.13 summarizes the differences and the correspondences between the bridging models and our approach.

In the following sections, we will present a performance comparing model for speed skating that strongly related to the bridging approach of Berry et al., but deviates on the points, globally described in Table 2.13. In case of speed skating, we do not have available a standard seasonal performance score (such as the number of goals in one season). In stead we use the r- times of the skaters. Because the races are organized on different rinks, the r-times are not only influenced by the season but also by the (location of the) rink. As rinks have improved and changed over the years, they are season dependent. Hence, the relative difficulty $y_t$ in the above equation should be specified as the relative difficulty per rink $r$ per season $t$, .i.e., we should take $y_{rt}$ in stead of $y_t$.

Moreover, we will assume that the individual abilities are continually influenced by the maturity level, while in Berry et al. this influence is partitioned into non-overlapping decades. The maturity level influences both the mean and the variation of all individual performances, and are therefore being modeled as a multiplicative effect. Furthermore, it is assumed that the maturity effect can be represented as a nonlinear function over time. In terms of the notation from Berry *et al.* (1999), we

**Table 2.13.** Global difference between bridging and our approach

| Description | Berry et al. | Our approach | Section location |
|---|---|---|---|
| Correcting for technological innovations and seasonal influences. | Taken into account by season/year dummies, representing the average performance level of that season year. | 'Eliminated' by taking first differences. | Section 2.6.1 |
| Increasing maturity level. | Individual abilities of athletes depend on the decade they are born in. | Corrected by estimating a non-linear function, representing the increase in maturity level. | Section 2.6.4 |
| sect:corfac Performance score. | Seasonal score. | Distance race differences aggregated to a seasonal score. | Section 2.7.1 |
| Peak performance ranking score. | Performance at the peak age of the sport | Performance during a number (4 or 5) best seasons. | Section 2.7.2 |

could formulate our approach uses:

$$q_i(t) = m_t * q_i$$

with $m_t$ the maturity influence in year $t$. and $m_t$ a non- linear function in $t$. In Section 2.6.4, this model formulated using our own notation.

Our season score is calculated by the aggregation the individual performances scores of the various tournaments, taking into account the importance of these tournaments.

Berry *et al.* (1999) use a general aging function to model the rise the decline in performance during an athlete's career. It is assume that each sport has a certain shaped aging curve, that varies between individuals in steepness, but with a peak at a fixed age. Based on the performance in this peak age the athletes receives his ranking score.

Finally we mention here that, instead of modeling performances by means of aging functions, we use four or five of the seasons in which the skater has his best performance. The choice between four and five will be explained in Section 2.7.2. Skaters try to 'peak'at more than one moment during their careers, because the focus is on the Olympic Games, organized each four years.

## 2.6 The USS performance score

Absolute skating times are not suitable for comparing performances because they do not satisfy the PS criteria from Section 2.4.3. As seen in Section 2.3.1, absolute times are influenced by innovations and the maturity rate. In the following sections we will show that relative race times are better measures than the absolute times. The actual performance measure is calculated in two phases. First, a fixed time per tournament race is determined and all skating times are transformed by subtracting this value from the realized skating times. In the second step, the resulting differences are corrected for the influence of the maturity rate. It will be shown that both the means and the variations of the differences decrease over time due to the influence of the maturity rate. The resulted score is called the USS-performance score.

### 2.6.1 Box plots of r-times

We start with analyzing absolute race times which are presented as r-times; see Section 2.2.4. In Figure 2.8, for both men and women, all r-times of the complete dataset are depicted conditioned on the seasons by means of box plots. On the horizontal axis the seasons of the dataset are depicted and on the vertical axis the r-times. For each season, the median of the r-times is depicted by means of a small horizontal line in the box. The database $DS$ does not contain times of races in which the skater fell during his race. It may happen that during a race a skater falls, finishes, and clocks a time. These finishing times are not included in our $DS$.

In Figure 2.9, the interpretation of the box plot is explained in detail. The median, $M$, is depicted by a small vertical line in the box. The left side of the box represents the first quartile of the r-times, and the right side the third quartile. So, each box contains 50% of all r-times realized in the corresponding season. The length, $IR$, of the box is defined by the interquartile range between the first and third quartile. On both sides, the box is extended by two lines with square bracket ends; the interval on the left hand side covers the 25% best r-times of the season and the right hand side the 25% of the worst. Let $T$ be the set of all r-times in a season. Then the upperbound of the interval is determined by $\min\{\max_{t \in T}(t), M + 1.5IR\}$, and the lower bound by $\max\{\min_{t \in T}(t), M - 1.5IR\}$; see Figure 2.9. All r-times outside this interval are called outliers and are marked by a '+'. These outliers are either extreme fast, or extreme slow r-times.

The box plots from Figure 2.8 show as expected a decreasing trend. Notice that the length of the lines and the length of the boxes (interquartile range) decrease as well. The median of the r-times decreases from 60.00 seconds in 1896 to 37.04 seconds in 2007 for the men, and from 63.51 in 1947 to 40.36 seconds in 2007 for the women. The interquartile ranges show a similar decrease, namely from 9.80 to 2.70 for the men, and from 9.00 to 1.52 for the women.

In case all absolute r-times are projected on the vertical axes, creating a one-dimensional ranking, the list becomes highly time dependent (it becomes comparable with the current fastest time rankings and the Adelskalender): all 'old' r-times

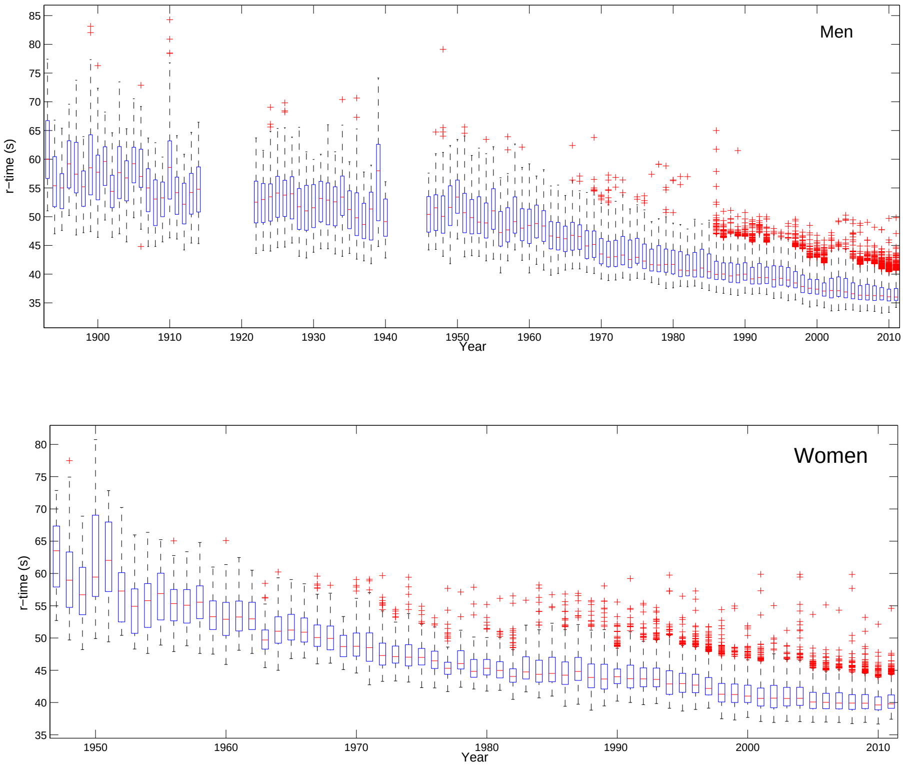



**Figure 2.8.** Box plots of all r-times

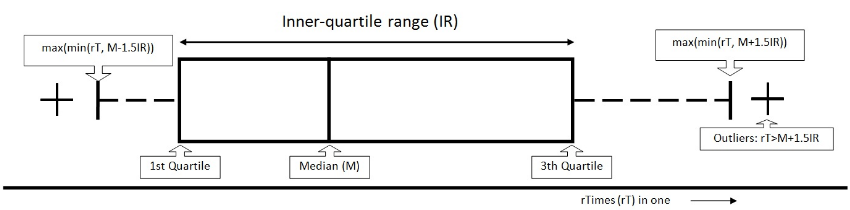



**Figure 2.9.** Box plot interpretation (rotated $90^o$)

are placed 'above' the recently realize r-times, meaning that all current skaters are 'better'. The absolute times conflict therefore with criterion PS 1 and PS 2 from Section 2.4.3. Hence, absolute times are not useful for comparing performances of skaters over long periods of time.

### 2.6.2 AV-values

The first step in constructing a performance score that satisfies all performance score criteria (see Section 2.4.3) is to eliminate the influence of technological innovations; see Section 2.3.2. Factors like rink and equipment are more or less constant within one tournament distance race and can therefore be considered as having an equal, or no, influence on the final order of a tournament distance race.

**Assumption 1.** *The influences of rink and equipment innovations on absolute times of one distance race are equal for all skaters of that race.*

The performance should therefore not be related to absolute times but to relative times within one tournament distance race. In Section 2.6.2, we will show that under Assumption 1, relative times within a tournament eliminate the rink and equipment effects described in Section 2.3.2.
The basis for our performance score will be the difference between the r-time and the average r-time of the first five skaters of a distance race [4]. This means that for each distance, the average r-time of the first five skaters is subtracted from the r-times of all skaters in that race. The resulting value is called the *AV5-value* of a skater.

Clearly, the first two or three skaters (depending on the average value) of a race will get a negative AV5-value, and all others a positive value. The value measures how much faster/slower the skater is in comparison to the average of the best five. All race winners will get the lowest value, and the lower this (negative) value is, the better the performance.

In Section 2.6.2, the exact calculations of the AV5-value are given, and we will discuss why the first five skaters are used. In this section, it is also shown that under certain assumptions the AV5-values are independent of technological innovations.

#### Calculation of AV$\alpha$-values

In the previous subsubsection an AV5-value is defined as the difference between the r-time and the average r-time of the first five skaters in one distance race. For any $\alpha > 0$, the AV$\alpha$-value of a skater in a certain distance race is the difference between the r-time of that skater and the average r-time of the first $\alpha$ skaters of that race. These values are calculated in the following way.

For $i \in S$ (set of skaters), $t \in Y$ (set of seasons), $k \in K$ (set of tournaments), $d \in D$ (set of distances), and $\alpha \geq 1$, we define,

$$I^{\alpha}_{tkd} \quad = \quad \{i \in S | 1 \leq \alpha \leq R_{itkd}\},$$

i.e., $I^{\alpha}_{kdt}$ is the set of skaters who finished within the first $\alpha$ positions on distance $d$ at tournament $k$ in season $t$. The average of the r-times of these first $\alpha$ skaters is denoted and defined as

$$\overline{rT}^{\alpha}_{tkd} = \frac{1}{\alpha} \sum_{i \in I^{\alpha}_{tkd}} rT_{itkd}.$$

[4] This can be seen as a trimmed mean with fixed boundaries; see Welsh (1987)

The difference with the trimmed mean definition of Welsh (1987) is that both the lower bound and the upper bound are fixed to, respectively, the first and the fifth observation of the order statistic, instead of quantiles.

Now, for each $(i, t, k, d) \in DB$, and each $\alpha \geq 1$, define

$$AV^{\alpha}_{itkd} = rT_{itkd} - \overline{rT}^{\alpha}_{tkd}.$$

Hence, $AV^{\alpha}_{itkd}$ is the difference in time (seconds) between the average time of the first $\alpha$ skaters and the r-time of skater $i$ on distance $d$ of tournament $k$ in year $t$. In the following section we will show that, based on a number of assumptions, these AV$\alpha$-values measure the relative individual performances of skaters, and that they are not subjected to technological innovations.

**Elimination of the influence of technological innovations**

During one distance race, we have already assumed (Assumption 1) that improvements in rinks and equipment, simply summarized as technological innovations, have no substantial influence on the final ranking of that race. If we make the following additional assumptions, we will show that the relative differences are independent of the technological innovations.

**Assumption 2.** *The absolute r-time of a skater is the sum of his individual quality, the state of the technological development at that time and that rink, and a random factor.*

**Assumption 3.** *Changing circumstances within a tournament, such as weather and ice-condition, captured by the random factor, have no influence on the relative results.*

Under these assumptions, the r-time, $rT_{itkd}$, of skater $i$ at tournament $k$ on distance $d$ in season $t$ can be modeled as

$$rT_{itkd} = \mu_{tk} + q_i + \epsilon_{itkd},$$

with

$$\begin{aligned} \mu_{kt} &= \text{time part caused by the state of technology in season } t \text{ at tournament } k; \\ q_i &= \text{time part caused by the quality of skater } i; \\ \epsilon_{itkd} &= \text{random error term, with } \mathrm{E}(\epsilon) = 0. \end{aligned}$$

The random error term concerns influences that are not captured by skater's qualities or technological innovations. Examples are weather circumstances, ice conditions, bad luck, or influence of the opponent. The expected value of this random factor is assumed to be zero. So, on average, the r-time only depends on the quality of the skater and the influence of technological innovations.
Hence,

$$AV^{\alpha}_{itkd} \quad = \quad (q_i - \bar{q}_{tkd}) + (\epsilon_{itkd} - \bar{\epsilon}_{tkd}), \tag{2.1}$$

with

$$\begin{aligned}
\bar{q}_{tkd} &= \frac{1}{\alpha}\sum_{i\in I_{tkd}^{\alpha}} q_i, \quad \text{the average quality of the } \alpha \text{ best skaters,} \\
\bar{\epsilon}_{tkd} &= \frac{1}{\alpha}\sum_{i\in I_{tkd}^{\alpha}} \epsilon_i, \quad \text{the average error term.}
\end{aligned}$$

Note that in formula (2.1) the parameter $\mu_{kdt}$, the influence of the innovations, has disappeared. This implies that the AV$\alpha$-value of a skater only depends on his relative performance and the factor $(\epsilon_{itkd} - \bar{\epsilon}_{tkd})$. However, the latter factor is assumed to be small and does not have much influence on the performance and the final order of the distance race. Moreover, when averages of AV-values are taken (as we will do later on), the influence of the random factor converges to zero. So, on average, the value of $AV_{itkd}^{\alpha}$ only refers to the relative quality of the skater, and is not influenced by technology. Hence, the AV-values satisfies criterion PS 1 from Section 2.4.3.

**How to choose the value of $\alpha$?**

In theory, the value of $\alpha$ can be chosen between 1 and the number of participants of the race. If $\alpha = 1$, the performance is related to the time of the winner. In this case, all winners obtain the AV1-value 0, and all other AV1-values are times (in seconds) behind the winner. If $\alpha > 1$, winners will have negative values and the faster the winner is compared to the other ones, the lower his value. So, only if $\alpha > 1$, it is possible to distinguish between winners.

In case the value of $\alpha$ is chosen equal to the number of participants of the race, the AV$\alpha$-values will not be consistent between tournaments, because the number of participants is not always the same; see Section 2.3.5. The AV-values will become less comparable as, in general, winners of tournaments with many participators will get lower values, not because they skate faster, but the mean value of the r-times of more skaters tends to become larger. So in case AV-values are calculated with $\alpha = R_{itkd}$, they violate criterion PS 3.

However, choosing a fixed value for the complete period, 1892 - 2011, the participation level will still bias the AV-values. In Figure 2.4, we have seen that the total number of active skaters increases over the years. More skaters during a tournament increase the competition level, meaning that the performances of the best $\alpha$ skaters are closer together. Hence, in general , AV-values of top skaters are smaller if a race has more participants and the $\alpha$-value is fixed.

Figure 2.4 shows that from 1955 on, at least 24 participants took part at any of the two allround tournaments. Also note that three periods can be distinguished, namely 1892-1915, 1916-1955, and 1956-2011. In the period 1893-1914, the average participation number lies around 16, between 1915 and 1952 around 22, and in the last 65 years around 28. The ratio of these participation levels is roughly 3:4:5.

Since the first two periods have a lower participation level, less skaters are competing for the victory. A fixed value of $\alpha$ for all seasons will therefore result in the fact that the winners of the early two periods obtain too low values, as the quality of the best $\alpha$ is less high. This problem is solved by making $\alpha$ dependent on both the season $t$ in which the tournament is organized, and the number of participants $N_{tdk}$ in the distance race of tournament $k$.

So we let $\alpha$ depend on the season $t$, distance $d$, and tournament $k$, in the following way[5]

$$\alpha_{tdk} = \begin{cases} \min([0.2N_{tdk}], 3) & \text{if } t \leq 1955,\ d \neq 10000\text{m} \\ \min([0.4N_{tdk}], 3) & \text{if } t \leq 1955,\ d\text{=}10000\text{m} \\ 5 & \text{otherwise.} \end{cases}$$

For the period 1955-2011, when all tournaments have more than 24 participants, we take $\alpha = 5$, based on the following assumption.

**Assumption 4.** *From 1955 until now, most top skaters are present at the major tournaments and the five best skaters are representative for the quality of that period.*

Based on the average participation level in the period 1955-2011, the five best skaters are roughly 20% (28/5) of the participating skaters at allround tournaments. So choosing $\alpha$ equal to 20% of the participation level, means for the period 1893-1915, on average, $\alpha = (0.2)(16) \approx 3$, and for 1924-1955, $\alpha = (0.2)(22) \approx 4$. In this case, the AV-values of the winners in races with few participants are not overestimated.
Since on the 10000m, only twelve skaters are allowed to participate, we take 40% instead of the 20% in the period before 1955. For the women we always take the best five skaters, as they started in 1947 and participation levels where more stable.

Although $\alpha$ is not consistent, we will keep denoting $\alpha_{tdk} = \alpha$ and keep referring to the AV5-values.

### Example

In Table 2.14, the AV5-values of the first five skaters on the 1500m of the Olympic Winter Games 2002 are presented. So, $t = 2002$, $k = OG$, and $d = 1500m$ . The values of $AV^5_{itkd}$ are listed in the last column. In the third column the race times are given, and in the fourth column the corresponding r-times.

**Table 2.14.** AV5-values OG 2002, 1500m

| Position | Name | Race time (sec) | r-time (sec) | AV5 (sec) |
|---|---|---|---|---|
| 1 | Derek Parra | 1:43.95 | 34.65 | -0.32 |
| 2 | Jochem Uytdehaage | 1:44.57 | 34.86 | -0.11 |
| 3 | Ådne Søndrål | 1:45.26 | 35.09 | 0.12 |
| 4 | Joey Cheek | 1:45.34 | 35.12 | 0.15 |
| 5 | Ids Postma | 1:45.41 | 35.14 | 0.17 |
| Average of first five ($\overline{rT}^5_{tkd}$) | | | 34.97 | |

[5] where $[x]$ means rounding $x$ to the nearest integer.

**Box plots of AV5-values**

In Figure 2.10 the box plots of the men's AV5-values per season are plotted. In contrast to Figure 2.8, the boxes are now much more horizontally aligned. In fact all boxes lie between $0.0$ and $8.0$, and after 1920 the minimum is never below $-2.0$. We already showed that by taking the AV$\alpha$ values we eliminated the influence of technological innovations (criterion PS 1). Furthermore by relating $\alpha$ to the number of participants, we have taken into account the changing competition level (criterion PS 3). However, we not have corrected for the influence of the maturity level, and thereby criterion PS 2 is not satisfied. In the following part we show by means of the box plot that also the pre-condition 1 from Section 2.4.3 not yet holds. Figure 2.10

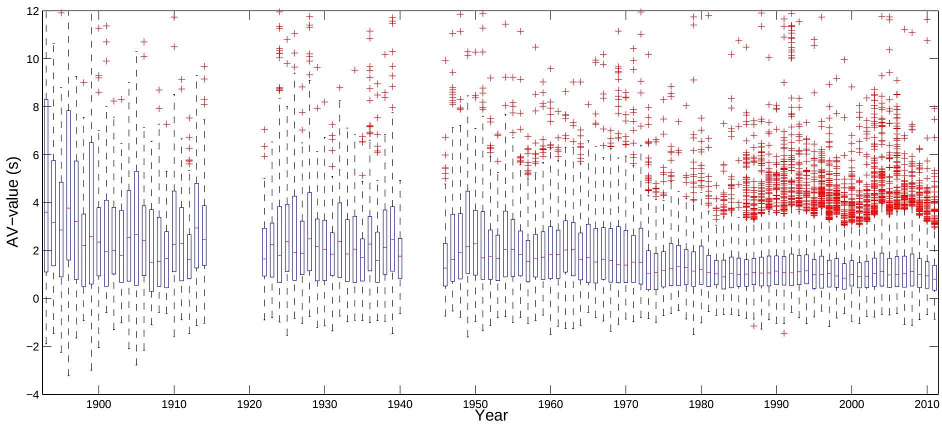



**Figure 2.10.** Boxplot of all AV5-values (Men)

shows that in the first ten years, the total range of the AV5-values is much larger than in the last twenty years. For example, in 1896 the minimum value is -3.32, while in 2004 the minimum is -0.66. The minimum values show a increasing trend whereas the maximum values show a decreasing trend.

Furthermore, the first and third quartile, represented by the end sides of the boxes, both converge to the median, indicating that the interquartile ranges (IRs) decrease over the years. The IRs of the first twenty-two years (1892-1914) are much larger than the IRs of last twenty year. Between 1892 and 1914, the first and third quartiles fluctuate within [0.5, 1.5] and [5.0, 8.0] respectively, resulting in an IR-value that varies within [5.0, 7.0]. In the period 1945-1980, the IR-values stabilize and drops to a value of around 3.0, where the first quartile values remain approximately equal to 1.0. Between 1980 and 2011, the first and third quartile are roughly 0.50 and 1.60, respectively, and so the IR-value lies around 1.10.

A consequence of using the AV5-values of Figure 2.10 as performance scores would be that the scores of the best skaters of first twenty years are better than the scores of the recent top athletes. Actually, the fastest skaters of the first twenty years would all be ranked higher, in contrast to what we have seen when absolute times are used (see Section 2.6.1).

Not only the scores of the best skaters of each year are incomparable, also most values from the IRs (on average within [1.5, 4]) of the period 1892-1940 are higher than the third quartiles (on average 2) of the period 1970-2010. In terms of performance scores, this would mean that average skaters from before 1940 perform worse than average skaters of the current generation. In the next section is explained why the medians and IRs of the AV-values still decrease, and how we can correct for this phenomenon.

### 2.6.3 Influence of the maturity level

The decrease of the median and interquartile range of the AV-values observed in Figure 2.10 can partly be explained by Gould's hypothesis; see Section 2.3.4. Gould states that complex systems, like sport, mature over extended periods of time, meaning that all performances will become 'better and the variance in performances decreases. As seen in Figure 2.10, based on the differences in r-times, it becomes more difficult for top skaters to distinguish themselves from average competitors. The minimum values and variations of the differences are smaller. Based on the minimum values of Figure 2.10 one may draw the conclusion that the current best skaters perform worse than the old champions. However, according to Gould a decreasing trend between the average and the top is not a surprise. He explains that the maturity effect causes that performances of average skaters become closer and closer to the top.

For the first 80 years of speed skating, Figure 2.10 confirms this hypothesis of Gould. In this period, the minimum of the AV-values converges to the median, meaning that top performances deviate less from average performances. It is also the case that both the average performances (the dots in the boxes), the third quartile and the maximum values decrease in this period, indicating an improvement in all performances. However, after 1972 the decrease is far less visible. Between 1971 and 1972, there is a decline in the third quartile, but after that period and especially after 1979, the length of the IR does not seem to decrease anymore. This may suggest that in the period 1972-2007 Gould's hypothesis does not hold.

#### Influence of participation level after 1972

The interquartile ranges of Figure 2.10 show a decreasing trend until 1972, and remain constant after that year. In order to explain this phenomenon, we distinguish, based on the length of the boxes (interquartile ranges), three periods, namely the periods 1892-1945, 1945-1972, and 1973-2011. In the first period and to a lesser extent also in the second period we observe a decreasing median and IR, confirming Gould's hypothesis (see Section 2.3.4). For the last period, however, after a downwards jump, the IRs remain at a constant level. All boxes have more or less the same length where, according to Gould, a decrease is expected.

The explanation for this fact is quite simple. At the start of the third period, namely in 1972, the World Sprint Championships are introduced, and in 1980 the first

World Cup Competition is organized. As shown in Figure 2.4, these tournaments increased the number of participants at international tournaments and, especially for the World Cups, this number has been growing ever since. Moreover, after 1972, sprinters exchanged the allround tournaments for the sprint tournaments, leaving open spots for other skaters to compete in allround tournaments. These developments allowed more skaters to compete, and therefore we observe performance data of a larger group of skaters. Therefore, in order to verify Gould's hypothesis correctly, we have to make a correction for the increasing number of participants.

In Figure 2.11, only box plots of the AV5-values of the fastest 24 (lowest restricted number on the World Allround Championships) skaters per tournament are plotted. Notice that the box plots of the period 1892-1914 remain unchanged. This is because the participation level never exceeded 24 (see Figure 2.4). For the period 1920-1938 the limit of 24 participants is exceeded only during a few seasons. Finally, Figure 2.11 shows that the box plots, especially after 1980, are smaller than the ones in Figure 2.10.

In Figure 2.11 both the median and the IR show a decreasing trend over the complete period. To confirm the differences between both figures, we have estimated the linear slope ($\beta_t$) over time of the median and the IR in three periods, namely 1892-2007, 1892-1972, 1973-2007 as follow. Let for each season $t$ and each $p$, $M_t^p$ and $IR_t^p$ be defined as

$$\begin{aligned} M_t^p &= \text{the median value of the } AV_{itkd}^5\text{-values in season } t \text{ with } R_{itkd} \leq p; \\ IR_t^p &= \text{inner quartile range of the } AV_{itkd}^5\text{-values in season } t \text{ with } R_{itkd} \leq p. \end{aligned}$$

Then for $p = max(N_{tdk})$ and $p = 24$, the following two linear models are estimated

$$M_t^p = \alpha_1 + \beta_1 t + \epsilon_{1t},$$

and

$$IR_t^p = \alpha_2 + \beta_2 t + \epsilon_{2t},$$

with $\epsilon_{1t}$ and $\epsilon_{2t}$ as normal distributed error terms.

Table 2.15 shows the estimated values of $\beta_1$ and $\beta_2$ of these linear regressions for both $p = max(N_{tdk})$ and $p = 24$. The table shows that all slopes are negative and that only the slopes of the median ($M$) and the inner quartile range ($IR$) for the period 1972-2011 are insignificant in case we use all skaters, i.e., $p = max(N_{tdk})$. In case only the best 24 skaters per tournament are considered the slopes remains significant negative for this period. This means that the increasing number of skaters and the introduction of new tournaments interfere with Gould's hypothesis.

The first conclusion, based on the results of Table 2.15, is that if only the best 24 skaters per tournament race are taken, the effect of Gould's hypothesis is visible,

**Table 2.15.** Estimation results

| | | Period (t) | | |
|---|---|---|---|---|
| **Variable** | **Number of positions** | 1892-2006 | 1892-1972 | 1973-2006 |
| $M_t^p$ | All | -0.013 (0.00) | -0.015 (0.00) | -0.002 (0.26) |
| | $\leq$24 | -0.017 (0.00) | -0.018 (0.00) | -0.012 (0.00) |
| $IR_t^p$ | All | -0.025 (0.00) | -0.025 (0.00) | -0.003(0.33) |
| | $\leq$24 | -0.030 (0.00) | -0.035 (0.00) | -0.013 (0.00) |

Estimation results of slope values $\beta_1$ and $\beta_2$ (P value) of linear regression between variable and seasons

i.e., the maturity level has influence on the average performance and its variation. Furthermore, we found that the correlation between the median and interquartile range in Figure 2.11 is 0.90, which means that they are strongly related and it looks that both the median and the IR decrease at the same rate. Hence, we make the following assumption

**Assumption 5.** *All AV-values are influenced by the maturity level by the same rate.*

So we may conclude that the AV5-values not satisfy PS 2 criterion of Section 2.4.3. In the following section is explained how the AV-values are corrected for the influence of the maturity level.

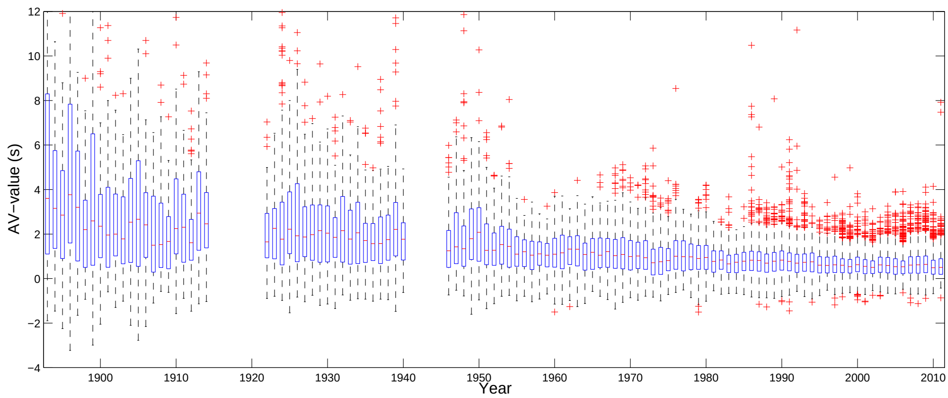



**Figure 2.11.** AV5-values of all top 24 ranked skaters (men)

### 2.6.4 Correcting for the influence of the maturity level

In this section it is shown how the effect of the maturity level on the AV5-values is eliminated, and how these values are transformed into a performance score that satisfies all criteria from Section 2.4.3.

#### The yearly average performance

In Section 2.3.1 we have introduced the concept of the General Performance Level (GPL), indicating an average level at which top skaters perform. We described that the GPL is partly influenced by the maturity level of the sport. In order to estimate the influence of the maturity level on the GPL of speed skating, we make the following assumptions

1. The relative change in the median of the AV5-values represent the relative change in the GPL and is completely caused by the maturity level.

2. The maturity level effects all AV5-values at the same rate.

The median of all AV5-values per season are taken as a gauge for the GPL, as it represents the performance of an average skater of that season. Since we already have corrected for technological innovations and other external factors, we may assume that the only influence on the GPL is the maturity level. The second assumption is based on the high correlation between the median and interquartile range found in Section 2.6.3. So we assume that the influence of the maturity level on the AV5-values can be calculated from the change of the GPL.

The GPL, represented by the median (see Figure 2.11), shows a decreasing (non)-linear trend and, as expected and assumed, a similar trend is seen in the top-side of the boxes (third quartile value, which in this case is proportional to the interquartile range, since the first quartile value is more or less constant). For example, comparing the GPL of 1900 and 2011, we observe that in 1900 the median value is 1.95, and in 2011 it is 0.54. Based on the assumptions made, it means that the maturity level has decreased the GPL by 72%. This decrease is comparable to the 75% decrease of the third quartile (from 3.26 to 0.81).

The influence of the maturity level between these two seasons can be eliminated by multiplying all AV-values of 1900 by the ratio 0.54/1.95. In this way all observations of 1900 are 'squeezed' and the median of 1900 is decreased to 0.54. Due to the high correlation between median and IR, this transformation will also decrease the IR of 1900 to the level of 2011.

We will apply this correction to all seasons using the season 2011 as a reference. After the correction, each season will have the same median as in 2011 and all AV-values are scores corrected for the influence of maturity. So, multiplying all AV5-values by the ratio of the median of the reference season (2011) and the median of the corresponding season, the resulting scores should be independent of the maturity level.

However, especially for the begin period, the median values fluctuate a lot between consecutive years. The correction values will therefore also fluctuate. In order to cope with these fluctuations, instead of using the real median values for the GPL, we model the GPL as a continues decreasing function over time. The median values per season are regressed as an non-linear function over time and the parameters of the non-linear function will be estimated. The estimated function values for the medians will be used for the correction, instead of the real median values.

**Estimating GPL's**

The GPL's are estimated for each discipline $L$; (see Section 2.4.1), since the disciplines use different sets of distances. Based on the conclusions from Section 2.6.3, we will only use the best 24 skaters of each distance race, resulting in adjusted median values of the AV5-values. Since during Allround tournaments a participation restriction is used for the male 10000m and the women 3000m (through 1983) and 5000m (after 1983), only the best 16 skaters for these distances races are used.

Let $t \in Y$, and $\alpha_{tdk} \geq 1$. Then the adjusted median AV5-value for men for each distance $d \in D_M$ is denoted and defined as

$$MAV^{\alpha}_{td} = \textbf{median} \bigcup_{k \in K} \{AV^{\alpha}_{itkd} \mid i \in I^{\eta_d}_{tkd}\}.$$

with

$$\eta_d = \begin{cases} 24 & d \in \{500m, 1000m, 1500m, 5000m\} \\ 16 & d = 10000m \end{cases}$$

For each discipline $L \in DS_M$, the adjusted median value is denoted and defined as

$$MAV^{\alpha}_{tL} = \begin{cases} MAV^{\alpha}_{td} & \text{if } L \in SD_M \\ \frac{1}{4} \sum\limits_{d \in L/\{1000m\}} MAV^{\alpha}_{td} & \text{if } L = OV_M \\ \frac{1}{2} \sum\limits_{d \in L} MAV^{\alpha}_{td} & \text{if } L = SP. \end{cases}$$

So, for $L = OV$ and $L = SP$, instead of taking the median of the whole set, we take the average of the median values over the distances of that set. Each distance will now have the same contribution on the medium over the seasons and the introduction of World Sprint Championships or World Cup Competition will not increase the influence of the short distances due to the increasing number of races. For the women we have *mutatis mutadis* the same definitions.

In order to illustrate the estimation process of the GPL, we have plotted in Figure 2.12 the values of $MAV^{\alpha}_{t,1500m}$. Especially in the first years, these values fluctuate heavily, meanly caused by the fact that the number of skaters is low in these years. When taking into account Gould's hypothesis, we have to model the MAV-values as a decreasing non-linear function. First, since the maturity level improves all performances over time, the medium values decrease.

Second, since the best skaters come closer to Gould's fictive boundary, the progression rate of the performances will slow down (non-linearly), and the performance level will converge to a certain boundary value (see Section 2.3.4). Since the medium values of the AV5-values can, by definition of AV5-values, never be smaller than zero, this boundary value is located somewhere between the current medium value and zero.

In Kuper and Sterken (2008a) and Grubb (1998) various nonlinear functions for fitting the development of world records over time are discussed. These functions

are based on biological grow models and are, although world records have a discrete jump nature, very useful to model athletic progress. In contrast to their research, we look at the progress of average performances of top athletes, instead of world records.
Three of the functions, namely the Exponential, the Weibull, and the Gompertz function, they use are tested on our data. The definition of these functions can be found in Table 2.16. In these functions, $\beta_1$ is the horizontal asymptote, and $\beta_3$ the grow parameter. The other parameters are extra fitting parameters. The parameters of these three functions are determined by fitting the adjusted median AV5-values as a GPL curve.

**Table 2.16.** Candidate GPL functions

| **Function** | |
|---|---|
| Exponential | $f(t,\beta) = \beta_1 + \beta_2 \exp^{\beta_3 t}$ |
| Weibull | $f(t,\beta) = \beta_1 + \beta_2 \exp^{\beta_3 t_4^{\beta}}$ |
| Gompertz | $f(t,\beta) = \beta_1 + \beta_2 \exp^{-\exp \beta_3 (t-\beta_4)}$ |

Now, for each discipline $L \in DS_M$, we model the MAV-values as

$$MAV_{tL}^{\alpha} = f(t, \beta_L) + \epsilon_{tL},$$

where $\beta_L$ is the optimal parameter vector, and $\epsilon_{tL} \sim N(0, \sigma^2)$ the residuals.

To estimate the $\beta$, we apply weighted nonlinear least squares[6], (see Dumouchel and O'Brien (1991)), an iterative process in which the weighted squared residuals are minimized. However, we use the number of observations per season as weights, since we want the best fit through seasonal medians with a large number of observations and being robust for outliers. In the first years of our data set, the medians are based on a relative low number of observations and so are less accurate. Since the medians of later periods are more accurate, we want a better fit through these points. This can be accomplished by multiplying the residuals with the number of observations. In this way, seasons with the highest number of observations get the best fit.

The estimated $\hat{\beta}$-parameters are used for the estimated performance level $G\hat{P}L_{tL}^{\alpha}$, given by

$$G\hat{PL}_{tL}^{\alpha} = Y(\hat{\beta_L}).$$

In Figure 2.12, for each of the three functions, the values of $G\hat{PL}_{t1500}^{\alpha}$ are presented. The figure shows that the three functions (Exponential, Weibull and Gompertz) are almost identical. Based on the final ranking results and the sensitivity analysis, the Gompertz function turns out to be the best fit for all disciplines. So all corrections are based on the values of the Gompertz function.

We may further note that for the women's 1500m and 3000m we needed to use an extra dummy variable for the seasons after 1983. In 1983, the 1000m in the allround tournaments is replaced by the 5000m. For the 1500m, this yielded that many

[6]Using the build in function nlinfit from MATLAB (2010)

short distance skaters, who before 1983 participated at the allround tournaments, no longer participate as the 5000m is too long for them. They were replaced by allround skaters. This change cause an upwards shift in the median values of the AV5-values of the 1500m in 1983, after which they decreased again. As the 3000m is no longer the longest distance, the participation restriction also shifted from the 3000m to the 5000m. This causing the same effect, despite the fact that we only look at medium values of the best 16.

In the next section, it is formulated how the values of $M\hat{A}V^{\alpha}_{tL}$ are used to correct the AV5-values for the influence of the maturity level.

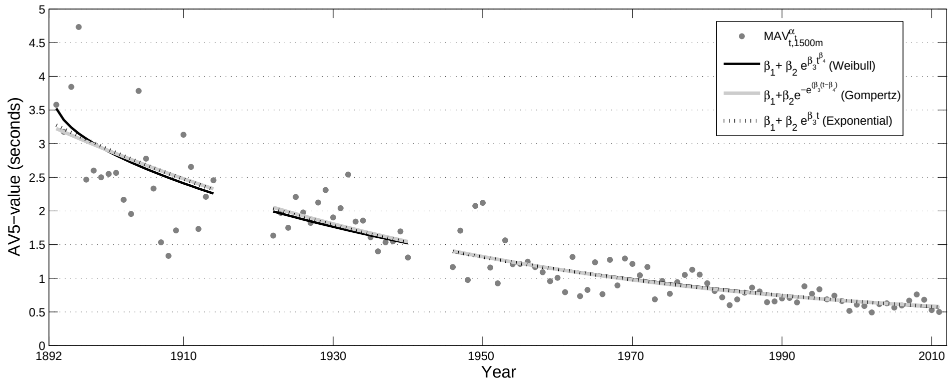



**Figure 2.12.** Estimation of the MAV-values, Men, 1500m

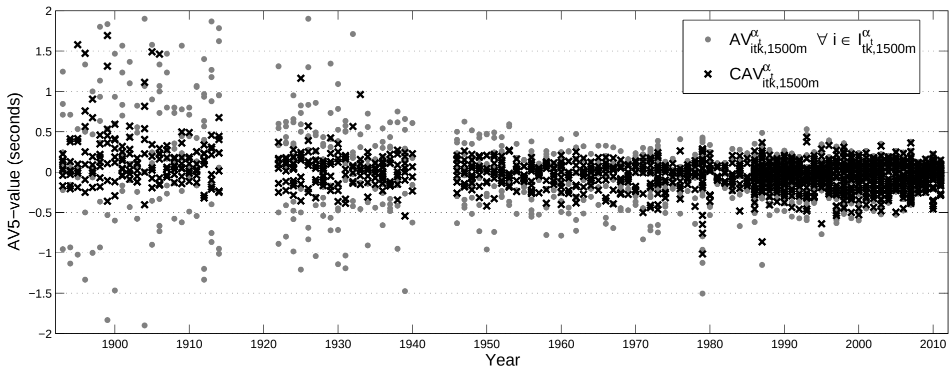



**Figure 2.13.** (C)AV-values, top 5 per distance race, Men, 1500m

## The correction factor

Now, as the estimated general performance level $G\hat{P}L^{\alpha}_{tL}$ of each season has been defined, the correction factor, denoted by $CF_{tL}$, will be introduced. The correction will be applied to all AV5-values and is supposed to eliminate the maturity effect on the GPL. This assertion will be tested in Section 2.6.5. For each season, the correction

factor is defined as the ratio between the $\hat{GPL}$ of that season and the $\hat{GPL}$ of a fixed season, in our case 2011.
So, for each season $t$ and discipline $L$, we define

$$CF_{tL} = \frac{\hat{GPL}^{\alpha}_{2011,L}}{\hat{GPL}^{\alpha}_{tL}}.$$

Recall from Section 2.6.4, that it is assumed that the increase in medium, representing the GPL, is completely caused by the maturity level and all AV5-values are influenced by the same rate. Based on that, we define and denote the corrected AV5-value by

$$CAV^{\alpha}_{itkdL} = CF_{tL} \times AV^{\alpha}_{itkd}.$$

In Figure 2.13, both the $AV^{\alpha}_{itkd}$ (denoted by the dots) as the $CAV^{\alpha}_{itkdL}$ (denoted by the 'x') of the first five ranked skaters of each 1500m are depicted. The figure shows that especially the values of the first twenty years are squeezed by the maturity correction procedure. In these years, the value of $\tilde{GPL}$ was higher than the value of $\hat{GPL}^{5}_{2011,1500m}$ (see Figure 2.12), which means that the GPL was not as good as in 2011. The difference between 'top' and 'average' skaters was much larger. The result of the correction is that 95% of the CAV-values of the first five ranked skaters now lies between -1.0 and 1.0, whereas before the correction the range of the AV5-values of the best five skaters was [-2; 2]; see Figure 2.13.

**Box plots of the maturity corrected AV5-values**

The effect of the corrections for the maturity level on all AV5-values is shown in Figure 2.14; in Figure 2.15 only the best 24 skaters per distance race are plotted.

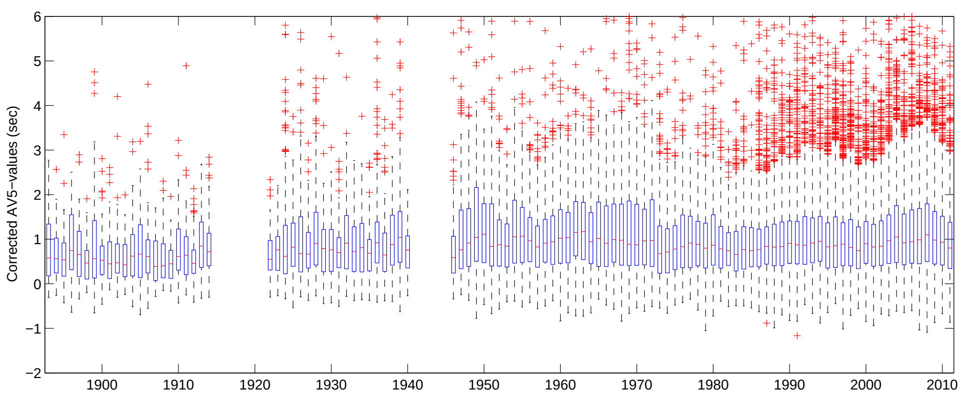



**Figure 2.14.** Corrected differences of all male skaters

Both figures show that all median values are now in the interval [0.5; 1], and that the decreasing trend has disappeared. The same holds, although to a lesser extent, for the first and third quartile and the corresponding IR. For the complete period,

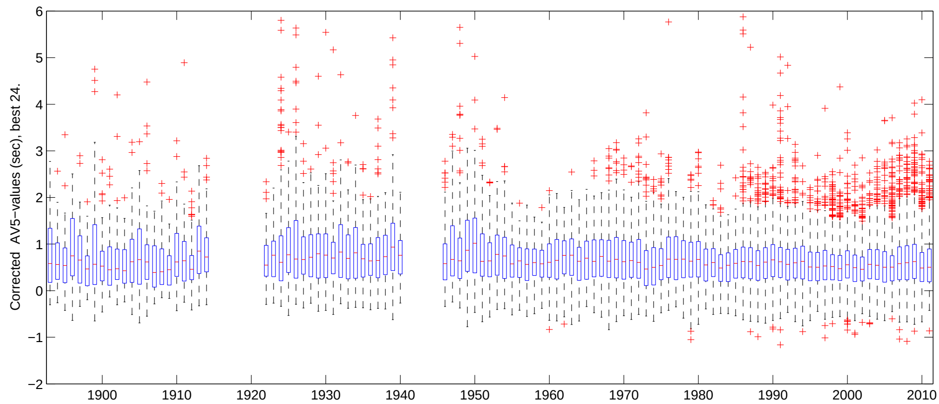



**Figure 2.15.** Corrected differences of all male skaters, only best 24

the first and third quartile values of Figure 2.14 fluctuate around the value of 0.35 and 1.46, respectively. The average IR is 1.10, and 83 of the 106 (78%) of these IRs lie within the interval [0.80; 1.40]. Recall that the large number of outliers in the last twenty-five years is a result of the introduction of the World Cups in which also skaters compete that do not belong to the absolute top.

### Example

The correction procedure for the individual AV5-values is illustrated by means of the following example. Consider the 1500m of the Olympics of 1972. Hence, $t = 1972$, $L = \{1500m\}$, $d = 1500m$, and $k = OG$. The first five skaters of this distance are listed in Table 2.17. In the second column the actual race times are given. The third column contains the r-times, and the fourth column the AV5-values. From Figure 2.12, one can derive that, $G\tilde{P}L^{5}_{\text{1972,1500m}} = 0.9379$ and $G\tilde{P}L^{5}_{\text{2011,1500m}} = 0.5841$. Hence, $CF_{td} = \frac{G\tilde{P}L^{5}_{\text{2011,1500m}}}{G\tilde{P}L^{5}_{\text{1972,1500m}}} = \frac{0.5841}{0.9379} = 0.6228$. So changing the AV5-values of the 1500m of 1972 to CAV5-values is achieved by multiplying all AV5-values by 0.6228. The last column of Table 2.17 shows the CAV5-values.

The CAV5-value -0.45 for Ard Schenk in Table 2.17 indicates that if this race would have been skated in 2011, his AV5-value would have been 0.45 seconds instead of 0.72 seconds. The CAV5-values can be used to compare results of other tournaments. In Table 2.18, we have listed the results and calculations of the Olympic 1500m in 1998. Comparing the results of both tables, it shows that, although the absolute times between both periods differs more than 15 seconds, the CAV-values are much more comparable. The ranges of the first five CAV-values in 1972 and 1998 ([-0.45; 0.30] and [-0.30; 0.36] respectively), have a large overlap, whereas the best absolute time of 1972 not even comes close the time of the number five of 1998. In case we compare all Olympic number one CAV-values it turns out that Ard Schenk's value of -0.45 is the best score ever made on an Olympic 1500m. In the following sec-

**Table 2.17.** Corrected AV-values, Olympic Winter Games 1972, 1500m

| Position | Name | Race time(sec) | Reduced 500m time | AV | CAV |
|---|---|---|---|---|---|
| 1 | Ard Schenk | 2.02,96 | 40.99 | -0.72 | -0.45 |
| 2 | Roar Gronvold | 2.04,26 | 41.42 | -0.29 | -0.18 |
| 3 | Goran Claeson | 2.05,89 | 41.96 | 0.25 | 0.16 |
| 4 | Bjorn Tveter | 2.05,94 | 41.98 | 0.27 | 0.17 |
| 5 | Jan Bols | 2.06,58 | 42.19 | 0.48 | 0.30 |
| Average of first five ($\overline{rT}^{5}_{tkd}$) | | | 41.71 | | |

**Table 2.18.** Corrected AV-values, Olympic Winter Games 1998, 1500m

| Position | Name | Race time(sec) | Reduced 500m time | AV | CAV |
|---|---|---|---|---|---|
| 1 | Adne Sondral | 1.47,87 | 35.96 | -0.33 | -0.30 |
| 2 | Ids Postma | 1.48,13 | 36.04 | -0.24 | -0.23 |
| 3 | Rintje Ritsma | 1.48,52 | 36.17 | -0.11 | -0.10 |
| 4 | Jan Bos | 1.49,75 | 36.58 | 0.30 | 0.27 |
| 5 | KC Boutiette | 1.50,04 | 36.68 | 0.39 | 0.36 |
| Average of first five ($\overline{rT}^{5}_{tkd}$) | | | 36.28 | | |

tion we will analyze in more detail the comparability of the CAV-values of different seasons but for the same distance race positions.

### 2.6.5 Testing the performance score criteria and the pre-condition

In Section 2.4 we have formulated three performance score criteria. The influence of technological innovations is eliminated by taking differences within tournaments (see Section 2.6.2). The influence of the maturity level is eliminated by estimating this effect on the GPL (see Section 2.6.4) and correcting all AV-values for this global effect (see Section 2.6.4).

Finally, by relating the number ($\alpha$) of best skaters to the number of participants and by using only the first 24 positions per tournament to estimate the effect of the maturity level, the influence of the changing number of skaters is taken in account.

In Section 2.4 we have formulated a pre-condition concerning performance scores for equal positions on different distance races. This pre-condition is tested in the following way. Take, for example, all skaters finished on the ninth place of the 500m of any Olympics. Both the AV5-values and CAV5-values of these skaters are plotted in Figure 2.16. The figure shows that the AV-values correlate (correlation of 0.81) strongly with the seasons, and that the CAV5-values are more or less season independent (correlation 0.01). The same conclusion can be drawn by analyzing the result of a simple linear regression through the points. The AV5-values have a significant negative slope ($\beta$), and the slope of the CAV5-values is zero. From this example, we may conclude that the CAV5-values, calculated for all Olympic ninth positions, are not season depended. So since these values do not depend on the season a skater was active in and thereby satisfy PS 2, we can use them for measuring the relative quality of a skater.

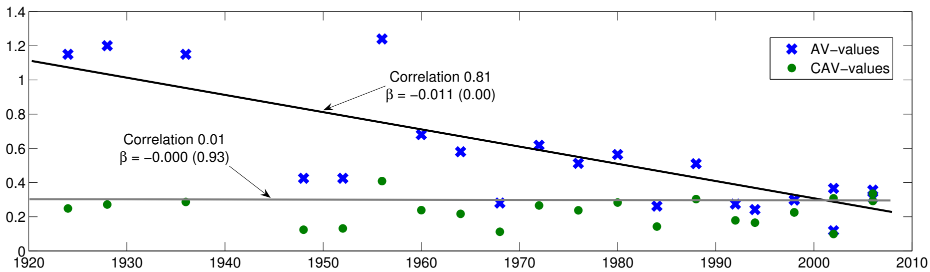



**Figure 2.16.** (C)AV5-values of skaters finished ninth on any Olympic 500m

Correlation value and slope as in example in Figure 2.16, are calculated for all tournaments and distances. In Table 2.19 a selection of the male results is presented, namely the 500m 1500m and 5000m of the Olympic Games, the 1000m of the WSCh, and the 1500m and 5000m of the WACh. The results for the WACh are based on the period 1945-2011, since before 1945 not all WACh for men have at least 24 participants.

**Table 2.19.** Correlations (C)AV-values and years for men

| | 500m OG | | | | | 1000m WSCh | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Position | 5 | 10 | 15 | 20 | 25 | 5 | 10 | 15 | 20 | 25 |
| Corr. AV | -0.52 | -0.63 | -0.86 | -0.90 | -0.84 | -0.47 | -0.58 | -0.64 | -0.67 | -0.74 |
| $\beta$ AV | -0.01 | -0.01 | -0.02 | -0.02 | -0.03 | -0.01 | -0.01 | -0.02 | -0.03 | -0.03 |
| t-prob AV | 0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Corr. CAV | 0.01 | 0.16 | 0.10 | 0.06 | -0.04 | -0.03 | 0.17 | 0.15 | 0.22 | 0.30 |
| $\beta$ CAV | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 |
| t-prob CAV | 0.98 | 0.52 | 0.68 | 0.82 | 0.88 | 0.82 | 0.15 | 0.20 | 0.06 | 0.01 |
| | **1500m OG** | | | | | **1500m WACh** | | | | |
| Position | 5 | 10 | 15 | 20 | 25 | 5 | 10 | 15 | 20 | 24 |
| Corr. AV | -0.59 | -0.63 | -0.82 | -0.81 | -0.73 | -0.48 | -0.65 | -0.72 | -0.66 | -0.71 |
| $\beta$ AV | -0.01 | -0.01 | -0.01 | -0.02 | -0.02 | 0.00 | -0.01 | -0.02 | -0.03 | -0.04 |
| t-prob AV | 0.01 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Corr. CAV | 0.04 | 0.30 | 0.12 | -0.02 | 0.00 | 0.16 | 0.22 | 0.09 | -0.06 | -0.06 |
| $\beta$ CAV | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| t-prob CAV | 0.87 | 0.24 | 0.64 | 0.95 | 0.99 | 0.18 | 0.06 | 0.48 | 0.61 | 0.65 |
| | **5000m OG** | | | | | **5000m WACh** | | | | |
| Position | 5 | 10 | 15 | 20 | 25 | 5 | 10 | 15 | 20 | 24 |
| Corr. AV | -0.55 | -0.68 | -0.65 | -0.58 | -0.57 | -0.50 | -0.58 | -0.66 | -0.69 | -0.74 |
| $\beta$ AV | -0.01 | -0.01 | -0.02 | -0.02 | -0.03 | -0.01 | -0.01 | -0.02 | -0.03 | -0.06 |
| t-prob AV | 0.02 | 0.00 | 0.01 | 0.01 | 0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Corr. CAV | -0.04 | 0.01 | 0.05 | 0.14 | 0.09 | -0.13 | 0.06 | -0.02 | -0.17 | -0.35 |
| $\beta$ CAV | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | -0.01 |
| t-prob CAV | 0.89 | 0.96 | 0.85 | 0.58 | 0.74 | 0.28 | 0.65 | 0.86 | 0.16 | 0.01 |

Corr. (C)AV= correlation between the years and the (C)AV-values. $\beta$ = regression coefficient between years and (C)AV-value. t-prob=t-probability of regression coefficient $\beta$.

The first three rows of each block in Figure 2.16 show the values of the correlations (Corr), the regression coefficient ($\beta$), and the t-probability (t-prob) of the AV5-

values. The other three rows show the statistics of the CAV5-values. The column refers to the position of the distance race that is compared. For example, the first row and first column shows that all AV-values of skaters finished fifth on the Olympics have a correlation of -0.52 with the years, and the regression coefficient has a significant value of -0.01.

The AV-values in each block of Table 2.19 have a high correlation with the seasons and a significant negative regression coefficient. Note that these facts support Gould's hypothesis. After the maturity correction, all correlations are lower than 0.40 and almost no regression coefficient is significant anymore, i.e., the CAV-values are time independent and all within the same range.

However, for the 25th position of the 1000m WSCh and the 24th position of the WACh there still is a small correlation and a significant negative slope. This is caused by the fact that a participation levels where not constant. In 1999 the WACh was restricted to 24 participators and due to the country and continent restrictions not always the best 24 where present.

For the 10000m, the AV5- and CAV5-values of equal race positions are somewhat harder to compare and not included in the table. This is because the number of skaters who are allowed to skate this distance has changed over the years. The CAV5-value of the number twelve in a field of 24 skaters is likely to be lower than the CAV5-value of the number twelve in a field of only 12 skaters, mainly because the country restriction causes the effect that if the number is reduced from twenty-four to twelve not always the worst twelve skaters are left out. If a country has more top 12 skaters than allowed, some of these top skaters have to stay home.

A second reason for not always having the best skaters of the world within the top 12 of the 10000m is that during an allround tournament good stayers may not qualify for 10000m. If their position in the ranking after three distances is not within the best twelve, or within the best eight of the 5000m, the are not allowed to start at the final 10000m.

In Section 2.6.2, it is explained how the AV5-values eliminate the influence of technological innovations and therefore satisfy criterion PS 1. In Section 2.6.3, we have eliminated the influence of the maturity level on the AV5-values by estimating the GPL in each season and correct for the differences. In this section, it is shown that the resulting CAV5-values are time independent and therefore satisfy criterion PS 2. In both corrections the changing participation numbers is also taken into account, so criterion PS 3 is also satisfied. Finally, in this section it is shown that equal distance race positions receive more or less the same CAV5-value, showing that pre-condition 1 made in Section 2.4 is satisfied. The conclusion is that CAV5-values can be used as performance scores. In Section 2.7.1 it is explained how these performance scores are used for our USS-ranking system.

## 2.7 The ranking model

In this section, the model is presented with which skaters will be ranked based on their best seasons, while the most important tournaments obtain the largest impact on the ranking. The CAV5-values of Section 2.6 are used to calculate scores for each season. Based on a number of these seasonal scores the skaters will be ranked.

### 2.7.1 Seasonal performance score

A CAV5-value, defined in Section 2.6.2, of a skater represents his relative performance on a certain distance race. However, as it is a score on one distance race, a CAV5-value still contains uncertainty in order to predict the relative quality of the skater, see equation 2.1. Mainly because during a race there are still small influences such as the behavior of the opponent, the weather, or the form of the day that may influence the observation. This uncertainty can be reduced by taking an average of a larger number of CAV5-values, for example, the average of a complete season. This average value indicates how well the skater has performed during that season, and is less influenced by uncertainty. Our ranking model will use such seasonal performance scores.

Actually, we take a weighted average value of all CAV5-values in a season, and the value of the weights are related to the type of tournaments and discipline in which the skaters will be ranked (see Section 2.7.1).

For each tournament $k$, the weight $w_{kL}$ is specified for three different discipline categories, namely the sprint category, ($L \in \{\{500m\}, \{1000m\}, SP)\}$, the individual distance category ($L \in \{\{1500m\}, \{3000m\}\{5000m\}, \{10000m\}\}$), and the overall category ($L = OV_{M/W}$) and we define

$$w_{kL} = \text{the weight of tournament } k \text{ when used in discipline } L.$$

For each $(i, t, d) \in DB$, we define

$$\begin{aligned} K_{itd} \quad &= \quad \text{set of tournaments in season } t \\ & \qquad \text{in which skater } i \text{ was active on distance } d. \end{aligned}$$

The seasonal performance score measured by the average weighted CAV5-value, indicating the performance of skater $i$ in season $t$ at discipline $L$, is calculated in the following way. For skater $i$, his seasonal performance score on discipline $L$ in season $t$ is denoted and defined as

$$P_{itL} = \frac{\sum\limits_{d \in L} \sum\limits_{k \in K_{itd}} w_{kL} CAV^{\alpha}_{itkdL}}{\sum\limits_{d \in L} \sum\limits_{k \in K_{itd}} w_{kL}}.$$

This average performance can also be calculated for an arbitrary set of seasons. Let $\Psi(\subseteq Y)$ be a set of seasons. We denote and define the average period performance score of skater $i$ on discipline $L$ during period $\Psi$ as

$$P_{i\Psi L} = \frac{\sum\limits_{t\in\Psi}\sum\limits_{d\in L}\sum\limits_{k\in K_{itd}} w_{kL} CAV^{\alpha}_{itkdL}}{\sum\limits_{t\in\Psi}\sum\limits_{d\in L}\sum\limits_{k\in K_{itd}} w_{kL}}.$$

Hence, the symbol $P_{i\Psi L}$ represents the performance of skater $i$ on discipline $L$ during the set of seasons $\Psi$. In Section 2.7.2, we present the model with which the best seasons from a skaters career are chosen. Based on these seasons, his best average performance is calculated.

### Tournament weight values

There are several reasons to distinguish between performances on different tournaments. First of all, not all tournaments are organized with the same frequency. Secondly, some tournaments are more importance than others. Skaters focus more on the important and prestigious tournaments, especially on the Olympic Winter Games; a victory at one of the Olympic distances is seen as more prestigious than winning a same distance at a WSDCh or a WACh. The distinction between tournaments is expressed by assigning weight values to the CAV5-values of the tournaments: The most important tournaments will get the highest weights. In Table 2.20, for each discipline $L$ and each tournament $k$, the weights ($w_{kL}$) are listed.

**Table 2.20.** Tournaments weight values $w_{kL}$

| | Discipline (L) | | |
|---|---|---|---|
| **Tournament** ($k$) | {500m}, {1000m} SP | {1500m}, {5000m} {10000m} | $OV_{M/W}$ $OV_{M/W}$ |
| Olympic Winter Games | 40 | 40 | 40 |
| World Sprint Championships | 20 | 0 | 10 |
| World Cup Competition | 2 | 1 | 1 |
| World Single Distance Championships | 20 | 10 | 10 |
| World Allround Championships | 0 | 10 | 20 |
| European Championships | 0 | 5 | 10 |

**Sprint category** The Olympic Winter Games are considered to be the most important for each discipline. For the sprint disciplines, 500m and 1000m, the World Sprint Championships and the World Single Distances Championships are considered to be of second highest importance. The World Cups are rated as 1/10 the World Sprint Championships weight. Each World Cup season, contains fourteen 500m races and ten 1000m races, which makes the total share of the World Cup Competition almost equal to halve share of a World Sprint Championships as 500m and 1000m are skated twice during this tournament. The results of 500m's of allround tournaments have no influence on the sprint scores.

**The individual distances** The third column of Table 2.20 shows that, for the 1500m, 5000m, and 10000m rankings, the Olympic Winter Games are weighted four times more than both the World Allround Championships and the World Single Distance Championships. World championships are less important than Olympic Games, while Olympic Games are organized only once in four years. For these three single distance disciplines, the European Championships are given half the weight of the World Allround Championships. Although the competition level of European championships is (slightly) lower than the other tournaments, the CAV-values for the best skaters do not, in general, differ much from the scores of the World Allround Championships. Furthermore, including the European Championships means more data, especially for skaters from the period in which only European and World Championships were held. Non-European skaters are not penalized as we take the average of all tournament scores in which the skater has participated. So different numbers of tournaments are allowed. This also makes it possible to include the new introduced tournaments. Again, the lowest weights are given to the World Cup Competition. The total weight of the World Cup Competition in one season is almost equal to the weight of the European championships as the 1500m, 5000m and 10000m are skated, respectively, 7, 5, and 2 times a season.

**The overall list** The overall category ($OV_{M/W}$) uses a different weight ratio. It includes all distances and concerns both sprint and long distances. During allround tournaments both short and long distances are skated, but at Olympics Games skaters choose specific distances. Long distance specialists usually choose the 5000m and 10000m, while allrounders with a good sprint will choose the 1000m and 1500m. In order to prevent that the scores in the overall lists depend too much on one single distance, we choose to weight the World Allround Championships only half the weight of the Olympic Games. In the allround tournaments all skaters have to skate at least a 500m and a 5000m, which means that we have scores for both short and long distances. The influence of the tournament weights is tested in Section 2.9.

### 2.7.2 Best seasons

In Section 2.7.1 we have discussed a possibility of determining the average performance of a skater during a set of seasons by taking the weighted average of the CAV5-values of this period. The resulting $P_{i\Psi L}$ score indicates how well skater $i$ has performed during the period $\Psi$ on discipline $L$.

In the final ranking model, skaters are ranked based on their performances in a chosen number of seasons. One could take the whole career of a skater, but this leads to complications. Many skaters need time to reach the top and will not perform at a top level during their entire career. Especially for skaters with long careers, the first and last seasons are usually not the best onces. For this reason only the best seasons are taken and skaters are not 'punished' for worse seasons at the end or beginning of their careers. We also choose to rank skaters on a fixed number of seasons so that the scores of all skaters are based on more are less the same number of tournaments.

The selection of the best seasons from a skater's career is discussed in Section 2.7.2. First we specify the actual number of seasons that will be chosen.

### Number of best seasons

Let the collection of seasons in which skater $i$ was active on at least one distance from $L$ be denoted and defined as

$$\Pi_{iL} = \{t \in Y \mid \bigcup_{d \in L} K_{itd} \neq \emptyset\}.$$

In order to prevent that seasons with only very low weights are chosen, we require that the skater has participated in either at least four World Cups or in one of the major tournaments with weight $\geq 4$. So, for each skater $i$ and discipline $L$, a season $t$ is a candidate for 'best' season if

$$\sum_{d \in L} \sum_{k \in K_{itd}} w_{kL} \geq 4. \tag{2.2}$$

The next question is the number of seasons that should be taken from a skater's career. Clearly, the number of seasons should be large enough so that sufficient observations are included. One season is clearly not enough, since especially skaters of the early days have only one or two observations per distance (see Section 2.3.5). We also want to avoid that a skater with one outstanding season dominates a skater who has performed well during several seasons. A third objection against one season is based on the fact that we will introduce the requirement that at least one of the seasons is an Olympic season, see paragraph Tournament requirements. This would make all other non Olympic seasons redundant.

On the other hand, if too many seasons are chosen, the average score may be too much influenced by the worse seasons. Most skaters remain at the top for a period of three or four seasons. Obviously, skaters should not be punished for the fact that they have had long careers. Moreover, the more seasons that are chosen, the fewer skaters can be ranked. This becomes clear when the distribution of male skaters career lengths is observed. In Table 2.21, for each distance, we present the distribution of career lengths where the career length of a skater is determined by the number of active season that satisfy restriction (2.2). The table shows that the distribution within each distance is almost identical. Roughly 40% of the skaters who participated in one of the major tournaments, never entered one of these tournaments for a second season and one fifth of the skaters was only active for two seasons. About 2% manage to participate in a major tournament 11 times or more.

Table 2.21 tells us that around 60% has less than three seasons with sufficient weight. Looking at the best performance during an WACh, we see that in general, these are not the best skaters. Furthermore, Figure 2.17 shows that skaters with only one active season are fairly uniformally distributed over the seasons. In each period of ten seasons around 50 skaters (five per season) debute but never return. These

**Table 2.21.** Distribution of male skater's career lengths.

| Distance ($d$) | Active seasons 1 | 2 | 3 | 4 | 5-10 | 11- | Total # skaters |
|---|---|---|---|---|---|---|---|
| 500m | 40% | 19% | 12% | 9% | 16% | 2% | 990 |
| 1000m | 36% | 17% | 12% | 8% | 21% | 3% | 584 |
| 1500m | 41% | 19% | 10% | 7% | 19% | 2% | 1483 |
| 5000m | 42% | 20% | 11% | 6% | 19% | 2% | 1410 |
| 10000m | 44% | 17% | 11% | 7% | 18% | 2% | 794 |

are usually rookies who get a chance for one season but don't make it to the top, or skaters who only participate at one Olympic Games.

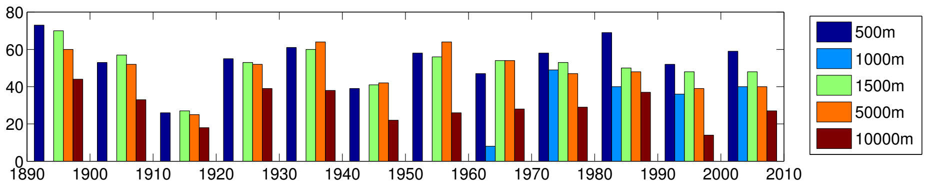



**Figure 2.17.** Distribution of skaters with an international career of one season

A ranking based on three seasons will contain 40% of the skaters and none of the important skaters is left out. Based on these arguments the following choice is made. For each $i \in S$ and $L \in DS_M$, let

$$\begin{aligned} \Psi_{iL}^{B} &= \text{the set of best seasons from} \\ &\quad\ \text{the career of } i \text{ on discipline } L \\ s_L &= |\Psi_{iL}^{B}|, \end{aligned}$$

and the value of $s_L$ is given by:

$$s_L = \begin{cases} 3 & \text{if} \quad L \in SD_M \\ 4 & \text{if} \quad L \in \{SP, OV_M\} \end{cases}$$

In Section 2.9.3, it is explained why for the overall list one extra season is chosen. The same values are used for the women disciplines.

### Tournament requirements

Besides the requirement that each season should have at least a total tournament weight of 4, we also make some special tournament requirements, namely:

1. At least **one** of the best season contains an Olympic distance race from $L$.

2. For the disciplines {500m}, {1000m}, and $SP$ skaters have to participate in at least **two** World Sprint Championships.

3. For the disciplines {1500m}, {5000m}, and {10000m} skaters have to participate in at least **two** either World Allround Championships or World Single Distances Championships.

4. For the discipline $\{OV_M\}$, it is required that the skater has participated in at least **twelve** world championship races (sum of WACh and WSDCh races) of which at least **two** are 500m races, and at least **two** are 5000m races.

The required tournament participation numbers will be denoted and defined as

$$V_{kL} \quad = \quad \text{the number of required participations at tournament } k \text{ in the best } s_L \text{ seasons at discipline } L.$$

and the required distance race requirements will be denoted and defined as

$$VD_{kLd} \quad = \quad \text{the number of required participations at distance } d \text{ of tournament } k \text{ in the best } s_L \text{ seasons at discipline } L.$$

The values of $V_{kL}$ and $V_{kLd}$ are listed in Table 2.22. Now, for each $i \in S$, $t \in Y$, $k \in K_{itd}$, and $d \in D$, define

$$n_{itkd} = \begin{cases} 1 & \text{if skater } i \text{ participated in season } t \text{ at tournament } k \text{ on distance } d \\ 0 & \text{otherwise.} \end{cases}$$

Then, for each $k \in \{OG, WSCh\}$, each $L \in DS_M$, and each $i \in S$, the set $\Psi^B_{iL}$ of best seasons should satisfy the following restriction

$$\sum_{d \in L} \sum_{t \in \Psi^B_{iL}} n_{itkd} \geq V_{kL} \tag{2.3}$$

and for the world championships ($k = WACh$ or $k = WSDCh$), we have that

$$\sum_{d \in L} \sum_{t \in \Psi^B_{iL}} n_{itWAChd} + n_{itWSDChd} \geq \frac{1}{2}(V_{WAChL} + V_{WSDChL}). \tag{2.4}$$

**Table 2.22.** Tournament and distance race participation requirements

| Discipline (L) | Tournament (k) OG | WACh WSDCh | WSCh | Distance (d) 500m | 5000m |
|---|---|---|---|---|---|
| 500m | 1 | 0 | 4 | 0 | 0 |
| 1000m | 1 | 0 | 4 | 0 | 0 |
| 1500m | 1 | 2 | 0 | 0 | 0 |
| 5000m | 1 | 2 | 0 | 0 | 0 |
| 10000m | 1 | 2 | 0 | 0 | 0 |
| Sprint | 1 | 0 | 8 | 0 | 0 |
| Overall (Total) | 1 | 12 | 0 | 2 | 2 |

Finally, for both $L = OV$ and $d \in \{500m, 5000m\}$, we have that

$$\sum_{t \in \Psi_{iL}} n_{itWAChd} + n_{itWSDChd} \geq \frac{1}{2}(VD_{WAChLd} + VD_{WSDChLd}). \tag{2.5}$$

Under these tournament and distance restrictions, the best seasons of each skater are chosen by means of a nonlinear binary optimization model.

### The selection model

The set $\Psi^B_{iL}$ of best seasons from the career of skater $i$ on discipline $L$ will be chosen in such a way that the weighted sum of the CAV5-values is minimal and the restrictions 2.2-2.5 are satisfied. The decision variable $x_{itL}$ is for each $i \in S$, $L \in DS$, and $t \in Y$ defined as

$$x_{itL} = \begin{cases} 1 & \text{if } t \in \Psi^B_{iL} \\ 0 & \text{otherwise.} \end{cases}$$

and the average weighted CAV-value, denoted by $BP_{iL}$, over the set of best seasons is determined by

$$BP_{iL} = \min_{x_{itL}} \frac{\sum\limits_{t \in \Pi_{iL}} \sum\limits_{d \in L} \sum\limits_{k \in K_{itd}} w_{kL} CAV^{\alpha}_{itkdL} x_{itL}}{\sum\limits_{t \in \Pi_{iL}} \sum\limits_{d \in L} \sum\limits_{k \in K_{itd}} w_{kL} x_{itL}}$$

s.t

$$\begin{aligned}
\sum_{t \in \Pi_{iL}} x_{itL} &= s_L & (2.6)\\
\sum_{d \in L} \sum_{k \in K_{itd}} w_{kL} x_{itL} &\geq 4x_{itL} \quad \forall t \in \Pi_{iL} & (2.7)\\
\sum_{d \in L} \sum_{t \in \Pi_{iL}} n_{itkd} x_{itL} &\geq V_{kL} \quad \forall k \in \{OG, WSCh\} & (2.8)\\
\sum_{d \in L} \sum_{t \in \Pi_{iL}} (n_{itWACHd} + n_{itWSDChd}) x_{itL} &\geq \frac{1}{2}(V_{WAchL} + V_{WSDChL}) & (2.9)\\
\sum_{t \in \Pi_{iL}} (n_{itWACHd} + n_{itWSDChd}) x_{itL} &\geq \frac{1}{2}(VD_{WAchLd} + VD_{WSDChLd}) & \\
& \qquad L = OV, d \in \{500m, 5000m\} & (2.10)\\
x_{itL} &\in \{0, 1\} \ \forall t \in \Pi_{iL}. & (2.11)
\end{aligned}$$

So, for each skater $i$, the seasons $t$ for which $x_{itL} = 1$ belong to his set $\Psi^B_{iL}$ of best seasons and his best score on discipline $L$ is given by $BP_{iL}$. Restriction (2.7) ensures that exactly $s_L$ seasons are chosen, and restriction (2.8) requires that each season of $\Psi^B_{iL}$ has at least a total tournament weight of 4. Restrictions (2.9), (2.10), and (2.10) ensure that skater $i$ has participated in the required number of tournaments and distance races.

### 2.7.3 Olympic Games restriction

Restriction (2.9) with $k = OG$ of Section 2.7.2 requires that at least one of the best seasons is a season in which the skater has participated at the Olympics. However, not all skaters have participated in an Olympic tournament, because either there were no Games organized (period 1892-1924), or they simply did not participate. These skaters will never satisfy restriction (2.9) for $k = OG$, and therefore the model will not find a solution and score for these skaters.

In order to overcome this problem (2.9) with $k = OG$ is neglected for these skaters and replaced by a penalty. The penalty consists of two parts, namely a fixed part and a variable part. The fixed part is only added to the value of $BP_{iL}$ if there was an opportunity to participate in Olympic Games. This means that skaters from before 1924 will receive no fixed penalty.

The variable part depends on the number and weight of the other tournaments the skater took part in. For instance, skaters who took part in all the major tournaments except for the Olympics will receive a lower penalty than the ones who only skated World Cup races. We will use the weights $w_{kL}$ of the tournaments in which the skater has participated in to determine the value of the variable part of the penalty. The lower this sum of the weights is the higher the penalty.

#### Specification of the penalty function

The fixed and variable part of the penalty depend on the period in which the skater was active:

- Category I: Skaters *only* active before 1924, the year of the first Olympic Winter Games;
- Category II: Skaters, *still* active after 1924.

Since skaters from Category I never have had the opportunity to participate in Olympic Winter Games, the fixed part of the penalty will be zero for this category. The total penalty value $p$ that is given to a skater who never participated in the Olympics is calculated as follow:

For each skater $i$ and discipline $L$, let $\Psi_{iL}^{B*}$ be the set of best seasons without the Olympic Games restriction (2.9) and let $c$ be the label of the categories, i.e., $c \in \{I, II\}$. For each $c$ and $L$, we define

$$
\begin{aligned}
\gamma_{cL} &= \text{penalty value of the fixed part of the penalty;} \\
\beta_{cL} &= \text{penalty value of the variable part of the penalty} \\
&\quad\ \text{given per 'missed' weight point;} \\
\eta_{iL} &= \text{the number of 'missed' weight points;} \\
Z_L &= \text{the required value of tournament weight points,}
\end{aligned}
$$

The penalty function $p(\Psi_{iL}^{B*})$ is defined and calculated as

$$p(\Psi_{iL}^{B*}) = \gamma_{cL} + \eta_{iL}\beta_{cL}$$

with

$$\eta_{iL} = \max\{0, Z_L - \sum_{t\in\Psi_{iL}^{B*}} \sum_{d\in L} \sum_{k\in K_{itd}} w_{kL}\}.$$

In the penalty function both $\gamma_{cL}$ and $\beta_{cL}$ are values measured in seconds; the value of $p(\Psi_{iL}^{B*})$ will be added to the optimal value $BP_{iL}$ of the skater. For each discipline $L$ and category $c$, the values of $\gamma_{cL}$ and $\beta_{cL}$ can be found in Table 2.23. The function shows that $\eta_{iL}$ and so the variable part depends on the total sum of the tournament weights ($w_{kL}$) of the tournaments in which the skater participated in his best seasons. For each weight point this total weight differs from value of $Z_L$, the skater will receive a penalty equal to the value of $\beta_{cL}$. However, in case the total sum of tournament weights in the set of best seasons is at least equal to $Z_L$, the skater has participated in sufficient many other important tournaments and only receives the fixed penalty value. In the following section we explain how the values for $Z_L$ are chosen. In Section 2.9.2 we discuss how the values of Table 2.23 are chosen and how the implementation of the penalty function has influenced the rankings.

**Table 2.23.** Penalty parameters

| Discipline | Penalty parameter | | | |
|---|---|---|---|---|
| | $\gamma_{cL}$ | | $\beta_{cL}$ | |
| | I | II | I | II |
| 500m | - | 0.05 | - | 0.005 |
| 1000m | - | 0.05 | - | 0.005 |
| 1500m | 0 | 0.1 | 0.015 | 0.01 |
| 5000m | 0 | 0.1 | 0.015 | 0.01 |
| 10000m | 0 | 0.1 | 0.015 | 0.01 |
| Sprint | - | 0 | - | 0.005 |
| Overall | 0 | 0.1 | 0.005 | 0.005 |

### The required tournament weights

A skater will only receive a fixed penalty in case his total tournament weight is at least equal to the required tournament weight $Z_L$, i.e., $Z_L$ is the minimum total tournament weight that is needed to make the variable part equal to zero. In Table 2.24 the values of $Z_L$are given. We take the value of $Z_L$ equal to the sum of the weights of the most important tournament of $L$ besides the Olympics.

For $L \in \{500m, 1000m, SP_M\}$, the World Sprint Championships are the most important tournaments besides the Olympics and if a skater participates $s_L$ times in this tournament, $\eta_{iL}$ will have the value 0. Because both distances are skated twice during this tournament, the value of $Z_L$ is twice (four times for $SP_M$) the value of the weight.

For $L = 1500m, 5000m, 10000m$, Table 2.24 shows that $Z_L$ is chosen to be equal three times ($s_L$=3) the sum of the weights of the two allround tournaments (WACh and ECh). As from 1996 on, these distances are also skated at the World Single Distances Championships, these skaters have the advantage that they can reduce the penalty with races from the World Champions Single Distances, since this tournament also has a weight of 10.

Also the overall list $L = OV$ uses the weights of both allround tournaments, but these weights are multiplied by 3, because at least three distances are skated during allround tournaments. The penalty is not increased if skaters miss the 10000m.

**Table 2.24.** Penalty parameters, $Z_L$ values

| Dis. $L$ | $Z_L$ |
|---|---|
| 500m | $s_{\text{500m}}(2w_{\text{WSCh},L})$ |
| 1000m | $s_{\text{1000m}}(2w_{\text{WSCh},L})$ |
| 1500m | $s_{\text{1500m}}(w_{\text{WACh},L} + w_{\text{ECh},L})$ |
| 5000m | $s_{\text{5000m}}(w_{\text{WACh},L} + w_{k\text{ECh},L})$ |
| 10000m | $s_{\text{10000m}}(w_{\text{WACh},L} + w_{k\text{ECh},L})$ |
| Sprint | $s_{\text{SP}}(4w_{\text{WSCh},L})$ |
| Overall | $s_{\text{OV}}(3w_{\text{WACh},L} + 3w_{k\text{ECh},L})$ |

**Example**

In order to illustrate the calculation of the penalty function, we consider a male skater from around 1995, who will be ranked on the 1500m. We assume that this skater has never participated in the Olympics, and in his best three years he competed only once in World Allround Championships and three times in the European Champoinships. Furthermore, he skated the 1500m eleven times at the World Cup Competition.

From Table 2.20, we know that $w_{\text{WACh},1500m} = 10$, $w_{\text{ECh},1500m} = 5$ and $w_{\text{WCC},1500m} = 1$. Hence, $Z_{\text{1500m}} = s_{\text{1500m}}(w_{\text{WACh},1500m} + w_{\text{ECh},1500m}) = 3((10) + (5)) = 45$. The total sum of the weights of his tournaments is equal to $w_{\text{WACh},1500m} + 3w_{\text{ECh},1500m} + 11w_{\text{WCC},1500m} = (1)(10) + (3)(5) + (11)(1) = 36$, so that $\eta_{i\text{1500m}} = 45 - 39 = 9$. The values $\alpha$ and $\beta$ are equal to 0.1 and 0.01, respectively (see Table 2.23) and the penalty is therefore equal to $0.1 + (9)(0.01) = 0.19$ (seconds). This value will be added to his $BP_{i1500m}$.

### 2.7.4 Ranking model

The ranking score of skater $i$ on discipline $L$ is denoted by $BF_{iL}$, and is calculated as follows. The number of Olympic participations is denoted and defined as

$$y_{iL} = \sum_{d \in L} \sum_{t \in \Pi_{iL}} n_{itOGd},$$

and

$$PD_{iL} = 1 - \frac{y_{iL}}{max\{1, y_{iL}\}}.$$

So $PD_{iL} = 1$ if skater $i$ never skated an Olympic distance of discipline $L$, and 0 else. The best score of skater $i$ on discipline $L$ is denoted and defined by

$$BF_{iL} = BS_{iL} + PD_{iL} \cdot p(\Psi^B_{iL}),$$

where $BS_{iL}$ is the optimal value in the model

$$BS_{iL} = \min_{x_{itL}} \frac{\sum\limits_{t \in \Pi_{iL}} \sum\limits_{d \in L} \sum\limits_{k \in K_{itd}} w_{kL} CAV^{\alpha}_{itkdL} x_{itL}}{\sum\limits_{t \in \Pi_{iL}} \sum\limits_{d \in L} \sum\limits_{k \in K_{itd}} w_{kL} x_{itL}}$$

s.t

$$\begin{aligned}
\sum_{t \in \Pi_{iL}} x_{itL} &= s_L && (2.12)\\
\sum_{d \in L} \sum_{k \in K_{itd}} w_{kL} x_{itL} &\geq 4x_{itL} \quad \text{for each } t \in \Pi_{iL} && (2.13)\\
\sum_{d \in L} \sum_{t \in \Pi_{iL}} n_{itkd} x_{itL} &\geq V_{kL} \quad \text{for} k = WSCh && (2.14)\\
\sum_{d \in L} \sum_{t \in \Pi_{iL}} (n_{itWACHd} + n_{itWSDChd}) x_{itL} &\geq \frac{1}{2}(V_{WAchL} + V_{WSDChL}) && (2.15)\\
\sum_{t \in \Pi_{iL}} (n_{itWACHd} + n_{itWSDChd}) x_{itL} &\geq \frac{1}{2}(VD_{WAchLd} + VD_{WSDChLd}) \\
& \qquad \text{for } L = OV, d \in \{500m, 5000m\} && (2.16)\\
\sum_{d \in L} \sum_{t \in \Pi_{iL}} n_{it\text{OG},d} x_{itL} + PD_{iL} &\geq V_{\text{OG},L} && (2.17)\\
x_{itL} &\in \{0,1\} \quad \text{for all } t \in \Pi_{iL}. && (2.18)
\end{aligned}$$

Based on the value of $BF_{iL}$ skater $i$ is ranked in increasing order on discipline $L$, which means that the lowest value is ranked highest.

# 2.8 Results

## 2.8.1 Model solving

The model given in Section 2.7.4 has to be solved for each skater and each discipline (seven in total). For the men this means it has to be solved around (2000)(7) =14000 times. Per skater, the number of sets of seasons that has to be evaluated is rather small. For example, a skater with a career length of eight years, skating the Olympic once, has a maximum ($\frac{7!}{3!4!}$ $=$)$35$ combinations of seasons (one Olympic season with three out of the other seven) for the discipline Overall and a maximum of ($\frac{7!}{2!5!}$ $=$)$21$ for the individual distance disciplines. Since the average career length in this dataset is four seasons, for many skaters only a small number of combinations has to be evaluated. So the model is solved by complete enumeration, i.e. checking all combinations, for skaters without a Olympic season. For skaters with at least one Olympic season only combinations that have one or more Olympic season are considers. The procedure for calculating for each skater and each relevant discipline his/her ranking score is formulated in Algorithm 1.

Algorithm 1 shows that first two sets of seasons are created. The set $O_i$ of seasons in which skater $i$ participated at the Olympic distances of discipline $L$ (line 3). All other seasons that satisfy restriction (2.2) are used in the set $R_i$ (line 4). If the total number of seasons is large enough (line 5), and the skater has at least one Olympic season (line 6), then each combination of two/three seasons in $R_i$ and one season from $O_i$ are used in the set $COM_{iL}$ (line 7). In case the skater has no Olympic season, the set $COM$ is filled with combinations of three/four seasons from $R_i$, while the penalty dummy is set to 1 (line 9 and 10). Then for each set of seasons in set $COM_{iL}$, the ranking score is calculated (line 15). In case all restrictions are satisfied (line 17, the algorithm checks if the score is better than the previous best score (line 19) and the score and set of best seasons is updated if this is true (line 20 and 22). Finally, the penalty is added to the best score in case the skater had no Olympic Season (line 26).

**Input:** The selection of discipline $L$, the set of skaters $(S)$
**for** $i = 1$ to $|S|$ **do**
  $O_{iL} \Leftarrow \{t \in \Pi_{iL} | \, OG \in \bigcup_{d \in L} K_{itd}\}$
  $R_{iL} \Leftarrow \{t \in \Pi_{iL} | \sum_{d \in L} \sum_{k \in K_{itd}} w_{kL} \geq 4\}$
  **if** $|O_{iL}| + |R_{iL}| \geq s_L$ **then**
    **if** $O_{iL} \neq \emptyset$ **then**
      $COM_{iL} \Leftarrow \{t \cup r | r \in 2^{R_{iL}}, |r| = s_L - 1, t \in O_{iL}, t \cap r = \emptyset\}$
    **else**
      $COM_{iL} \Leftarrow \{r | r \in 2^{R_{iL}}, |r| = s_L\}$
      $PD_{iL} \Leftarrow 1$
    **end if**

    $BS_{iL} \Leftarrow \infty$
    **for** $j = 1$ to $|COM_{iL}|$ **do**
      $bs \Leftarrow \frac{\sum_{j \in COM_{iL}} \sum_{d \in L} \sum_{k \in K_{itd}} w_{kL} CAV^{\alpha}_{itkdL}}{\sum_{t \in \Pi_{iL}} \sum_{d \in L} \sum_{k \in K_{itd}} w_{kL}}$

      **if** $\sum_{d \in L} \sum_{t \in COM_j} n_{itkd} \geq V_{kL}$ for $k = WSCh$
      **and** $\sum_{d \in L} \sum_{t \in \Pi_{iL}} (n_{itWAChd} + n_{itWSDChd}) x_{itL} \geq \frac{1}{2}(V_{WAChL} + V_{WSDChL})$
      **and** $\sum_{t \in \Pi_{iL}} (n_{itWAChd} + n_{itWSDChd}) x_{itL} \geq \frac{1}{2}(VD_{WAChLd} + VD_{WSDChLd})$
        for $L = OV, d \in \{500m, 5000m\}$ **then**

        **if** $bs < BS_{iL}$ **then**
          $x_{itL} \Leftarrow \begin{cases} 1 & if \quad t \in COM_j \\ 0 & else \end{cases}$
          $BS_{iL} \Leftarrow bs$
          $\Psi_{iL} \Leftarrow COM_j$
        **end if**
      **end if**
    **end for**
    $\overline{BF}_{iL} \Leftarrow BS_{iL} + PD_{iL} \cdot p(\Pi_{iL}, x_{itL})$
  **else**
    $\overline{BF}_{iL} \Leftarrow \infty$
  **end if**
**end for**

**Algorithm 1:** Ranking score calculation

### 2.8.2 Ranking lists

The model in Section 2.7.4 is used to rank skaters on seven disciplines. The position in the ranking is determined by the performance value $\overline{BF}_{iL}$. The skater with the lowest value is ranked as the number one. For both men and women, the first twenty skaters of every list is presented in Table A.1 through A.14 of the Appendix. The first column of each table is the ranking position. The second column contains the name of the skater, and the third column the value of the performance score ($\overline{BF}$). In column four through seven the seasons on which the skater is ranked are given. The seasons are sorted in such a way that the season with the lowest performance score is listed first. The bold typed seasons are Olympic seasons; the names of the Olympic champions are also typed bold. Below, we shortly discuss the results of the first three skaters of each list.

**Men, 500m** The number one of the 500m is Uwe-Jens Mey. He won the 500m two times on the Olympics (1988 and 1992) and was five times the winner of the 500m of a World Sprint Championships. Hiroyasu Shimizu is ranked second, he won Olympic gold on the 500m in 1998, and won five times the final ranking (combination of two 500m's) on the 500m during the World Single Distance Championships. Eric Heiden, who won the Olympic 500m in 1980, is ranked third; he also became four times world sprint champion.

**Women, 500m** The number one is the current world record holder, Jenny Wolf. Although she became second during the OG of 2010, she was unbeatable in the period 2007-2010. Three times Olympic winner of the 500m (1988, 1992 and 1994), Bonnie Blair, is ranked on the second place; she also won eleven times the 500m on the Sprint Championships. The number three, Catriona LeMay, won twice the Olympic gold medal (1998 and 2002).

**Men, 1000m** The 1000m list for the men consists only of skaters from after 1970. Before this year no World Sprint Championships were organized, and the first Olympic 1000m was skated in 1976. The best 1000m skater is Eric Heiden, who won all 1000m's in his best three years (1978, 1979, 1980). The number two, Igor Zhelezovski, never won the Olympic 1000m, but between 1989 and 1993 he won most of the 1000m's during the World Sprint Championships and World Cup Competition. On the third place is ranked Gaetan Boucher. He won three times the 1000m on World Sprint Championships, and also won a golden Olympic medal on this distance.

**Women, 1000m** The 1000m for the women was until 1983 part of the allround championships. However, this ranking is mainly based on the performance of skaters on sprint and Olympic tournaments. The best female 1000m skater is Karin Enke. In 1984, she won the Olympic gold medal and between 1980 and 1987 she always finished within the first three on the 1000m. The number two, Natalya Petrusyova, dominated the 1000m in the period before Karin Enke. In 1980 she won the Olympic 1000m. Christine Nesbitt is third on the list. She won the Olympic 1000m in 2010 and was three times the world champion single distances.

**Men, Sprint** Eric Heiden is absolutely the best sprinter. Between 1977 and 1980 he won all four World Sprint Championships, plus the Olympic golden medals on the 500m and the 1000m. Igor Zhelezoski, number two, won also all World Sprint Championships in his best four years but performed worse on the Olympics. Uwe-Jens Mey did the opposite, he was successful on the Olympic but never won a World Sprint Championship.

**Women, Sprint** Karin Enke is the best sprinter of the women and almost achieved Eric Heiden's performance. Enke won four times the World Sprint Championships but only one Olympic goldmedal, namely on the 1000m. She finished second on the Olympic 500m of 1984. Bonnie Blair won the sprint tournaments of 1989 and 1994, and both distances on the Olympics of 1994. The number three, Natalya Petrusyova won the World Sprint Championships in 1982, Olympic gold on the 1000m and Olympic bronze on the 500m in 1980.

**Men, 1500m** The number one of the 1500m for men is Ard Schenk. Schenk won, with exception of the 1500m on the European Champoinships of 1971, all the 1500m's in his best three years. The number two of the list, Eric Heiden, won the Olympic title, and two times the 1500m on the World Allround Championships. In his third best season he finish second. Clas Thunberg is the number three of the list. In 1924 and in 1928 he won the Olympic 1500m, and in the period 1924-1931 never finished outside the top 3.

**Women, 1500m** Besides being an excellent sprinter,Karin Enke also was an excellent 1500m skater. She is ranked first on the 1500m before Cindy Klassen and Anni Friesinger. Enke won her Olympic title in 1984, Klassen in 2006, and Friesinger in 2002. Friesinger also won the distances six times during the World Champions Single Distances.

**Men, 5000m** Gianni Romme, the number one of the 5000m, was one of the first long distance specialists. He focused completely on 5000m and 10000m. In the first three years of his career, Romme only skated the 5000m on single distance tournaments. In in 1998 he won Olympic gold. Later on, he also participated in allround tournaments in which he also won 5000m's. Hjalmar Andersen, the second best 5000m skater of all times, won all 5000m's in 1950, 1951, and 1952. The number three of the list is

Johann Olav Koss. He won most of the 5000m between 1991 and 1995. His greatest victory was the Olympic gold medal of 1994 in Hamar.

**Women, 3000m** The first ranked skater is Gunda Niemann, who won the Olympic 3000m in 1992 and 1998. In these years she won most of the 3000m's during allround and single distances tournaments. Andrea Mitscherlich was the best long distance skater of the eighties and is ranked second. On the third place we find Martina Sablikova who became Olympic champion in 2010, and won the Single Distances Championships in 2007

**Men, 10000m** The first three of the 10000m are equal to those of the 5000m. Koss is the number one, mainly due to his superior race in 1994 at the Olympics, where he beat the number two with more than 18 seconds. The number two, Gainni Romme, had in 1998 an extraordinary Olympic race; he beat the number two with 10.43 seconds. Hjalmar Anderson is third on this list.

**Women, 5000m** The best 5000m skater is Gunda Niemann. She won the Olympic 5000m once, but made her best performances during allround championships. During the Olympics of 1998 she finish second 0.04 second behind the number three of the list, Claudia Pechstein. Pechstein won three times the Olympic gold medal but during all other tournaments, she finished behind Niemann. Second of the list is Martina Sablikova, who won Olympic gold in 2010 and all other 5000m's between 2007 and 2011.

**Men, Overall** Best skater of all times is Eric Heiden. Heiden won all five distances during the Olympic Games of 1980. From 1977 through 1979 he won all World Sprint and Allround Championships. After the Olympics of 1980, he became second during the World Allround Championships. The number two of the list, Ard Schenk, did almost the same. During the Olympics of 1972, he won three of the four distances and in his best four years he was four times European Allround Champion, and three times World Allround Champion. The number three, Sven Kramer, won only one Olympic title. His allround victories are more impressive: he won five world and five European allround titles.

**Women, Overall** Both Gunda Niemann, the best female skater of all times, and the number two, Karin Enke, won eight Olympic medals. However, Niemann managed to win more allround titles. In her four best years, she won all World Allround Championships. In total she won this tournament eight times. Enke succeeded only four times. The number three of the list, Lidia Skoblikova, won the World Allround Championships twice, and six Olympic gold medals during the Olympic Games of 1960 and 1964.

### 2.8.3 Performance and ranking score criteria

In Section 2.4.1 and 2.4.4 we have formulated criteria for the performance score and the USS ranking model. The performance score is supposed to be independent of technological innovations, maturity level, and the active skating population. In Section 2.6.5, we have shown and explained that the CAV5-values, used as performance scores satisfy these performance score criteria.
The ranking model, described in Section 4.8, uses the CAV5-values. The solutions to the model should satisfy criteria RS 1, RS 2, and RS 3. Criterion RS 1 states that the importance of a tournaments should have influence on the ranking score. In our ranking model the CAV5-values of the various tournaments are weighted in the average season scores. CAV5-values of the most important tournaments, such as the Olympic Winter Games, are weighted higher and therefore these scores have the highest influence on the ranking score.

Criterion RS 2 demands that results of newly introduced tournaments are taken into account. Recall that the weight models of Section 2.5.3 do not satisfy RS 2. Since the USS model uses a weighted average score over the tournaments, a higher number of tournaments will not automatically increase the ranking score. Moreover, the European Championships hardly give any benefit to the European skaters, because they still need to perform well at the World Championships such as to keep their average ranking score at the same level.

Finally criterion RS 3 is satisfied since the influence of the length of a career is eliminated by only looking at the best seasons of the skaters. Although skaters with longer careers have the advantage that they have more seasons to choose from, they need three or four seasons with at least two World Allround Championships or World Single Distance Championship races and one Olympic Game race.

In the following section we will analyze whether or not the results of the USS ranking model satisfy criteria RO 1 and RO 2 from Section 2.4.4.

### 2.8.4 Analyzing the output criteria

In this section the output criteria from Section 2.4.4 are reviewed. Conclusions on the quality of the ranking lists with respect to the output criteria are hard to draw, mainly because the output criteria itself are not formulated as measurable indicators. Our output criteria mainly serve as benchmarks for quality discussion. In the first part of Section 2.8.4 the ranking position of the Olympic champions is discussed. The obvious expectation is that they are in the top of the lists. In the second part of Section , we analyze whether the highest ranked skaters are more or less uniformly distributed over the years. Finally, in the third part of Section a sub-ranking of Dutch skaters is analyzed and we test if it is consistent with the performances the skaters.

#### Olympic Champions

In Section 2.5 we presented the skaters with the most tournament victories and the most world records. Criterion OR 1 yields that we expect them in the top of the

USS-rankings. If Table 2.3 is compared with the ranking of the Overall discipline, we observe that almost all skaters with three or more allround world titles are present in the top 10 of the Overall list. Only Michael Staksrud not present there, he is ranked 18th in the men's Overall. Erben Wennemars (16) and Akira Kuriowa (14) are the only skaters with more than two sprint titles, but not in the top 10 of the sprint ranking.

All Olympic Champions within the top 20 are typed bold in Tables A.1 through A.14. In Table 2.25 we list the percentage of Olympic champions that are either in the top 20, top 50, or not ranked. The Olympic Champions that are not ranked do not satisfy the tournament restrictions.

**Table 2.25.** Ranking position of Olympic champions

| **Distance** | **Top 20** | | **Top 50** | | **Not possible to rank** | |
|---|---|---|---|---|---|---|
| | Men | Women | Men | Women | Men | Women |
| 500m | 50% (9) | 81% (9) | 66% (12) | 81% (9) | 33% (6) | 19% (2) |
| 1000m | 70% (7) | 66% (8) | 90% (9) | 75% (9) | 10% (1) | 25% (3) |
| 1500m | 52% (9) | 92% (11) | 100% (17) | 100% (12) | 0% (0) | 0% (0) |
| 3000m | | 82% (9) | | 100% (11) | | 0% (0) |
| 5000m | 83% (15) | 100% (6) | 94% (17) | 100% (6) | 6% (1) | 0% (0) |
| 10000m | 68% (13) | | 89% (17) | | 6% (1) | |

**Men** Table 2.25 shows that on the 500m list nine of the eighteen Olympic winners are within the first twenty positions. The fact that 50% of the champions is absent can be explained as follow. The Olympic winners of 2002 and 2006, Casey FitzRandolph (39) and Joey Cheek (31), and the Olympic winner of 1994, Golubyov (46), have focused completely on the Olympics and did not perform well during the World Sprint Championships. However, they are still ranked within the top 50. The first five winners, namely Charles Jewtraw, Bernt Evensen, Jack Shea, Ivar Ballangrud, Finn Helgesen, and the winner of 1964, Terry McDermont, have no other results than this golden medal, since no sprint championships were organized before 1971.

In the top 20 of the 1000m seven of the ten Olympic champions are present. Two Olympic winners, Nikolay Guljajev(21) and Olaf Zinke (47), have relatively bad sprint championships performances, and the winner of 1998, Ids Postma, participated only once in the World Sprint Championships.

Nine of the seventeen Olympic 1500m champions are ranked within the best twenty. The other eight, Gaetan Boucher (21), Sverre Farstad (24), Ants Antson (26), Charles Mathiesen (27), Enrico Fabris (30), Derrek Parra (31), Hjalmar Andersen (36), Andre Hoffmann (43), are ranked within the first fifty.

For the 5000m almost all Olympic champions are present within the first twenty. Only three of the eighteen are not, namely Tomas Gustafson (30), Reidar Liaklev (33), and Irving Jaffee. The last one is not ranked at all, since no other results were found for this skater.

In the top twenty of the 10000m, thirteen of the nineteen winners are present. Again, Jaffee is not ranked, because no other results were found. The other five Sigge Ericsson (22), Igor Malkov (27), Julius Skutnabb (34), Ake Seyffarth(38), Johnny

Hoglin (55) performed excellent on the Olympics but never repeated that performance during other tournaments.

For the male skaters we can conclude that at least 50% of the Olympic champions can be found within the top 20. It can also be seen that the top 3 of the 500m, the 1500m, the 5000m, and the 10000m consists of Olympic champions. The remaining ranked Olympic champions are all in the top 50.

**Women** The women skaters were active thirteen times during Olympic Games in the period 1955-2011. Nine of the eleven 500m winners are ranked within the first twenty skaters. The first two winners, Helga Haase and Lidia Skoblikova are not ranked, because only skaters with at least one participation at the World Sprint Championships are ranked and this tournament was not organized before 1970.

In the 1000m ranking eight of the twelve Olympic champions are ranked. Due to the same reason as above for 500m, the first three winners Klara Goeseva, Carry Geijssen and Lidia Skoblikova, are not ranked.

Only one of the twelve Olympic 1500m winners is not ranked within the first twenty skaters, namely the Dutch sprinter Marianne Timmer, who surprisingly beat all the allround skaters during the Olympics of 1998: She never repeated this success during World Cups or Single Distance Championships.

At the 3000m even nine of the eleven are ranked within the first ten skaters. The other two, Svetlana Bazhanova and Tatyana Averina, are ranked on the 29th and 30th position.

The Olympic 5000m for women is only held six times, and all five winners (Claudia Pechstein won three times) are ranked within the first ten. Also all number two are within this first ten.

For the female skaters we found that at least 66% of the Olympic champions are ranked within the top 20. For the 1500m, the 3000m, and the 5000m all Olympic champions are in the top 50.

#### Distribution of top skaters over the years

Criterion RO 2 (see Section 2.4.4) stated that top skaters from the ranking lists should be uniformly distributed over the years. Each decade has skaters who dominate and win most of the prices. The CAV5-values satisfy pre-condition 1 and so winners will get on average the same score. Therefore, we may expect that 'decade' champions are in the top of the lists.

To verify this, histograms of the top 20 (black bars), top 50 (grey bars), and top 100 (white bars) of the seven disciplines are plotted; see Figure 2.18 and Figure 2.19. The histograms show the percentages of skaters from each decade, present in the top 20, top 50 and top 100. The horizontal axis refers to the decades. On the vertical axis the percentage are given (0.2 refers to 20%). The year used for a skater is the season in which a skater has obtained his best score. For example, the first black bar of the 500m, men histogram shows that 35% (seven skaters) of the top 20 had their best season in the decade 1970-1980. For the top 50 and top 100 this decade is presented for 30% and 24% of the 500m, respectively.

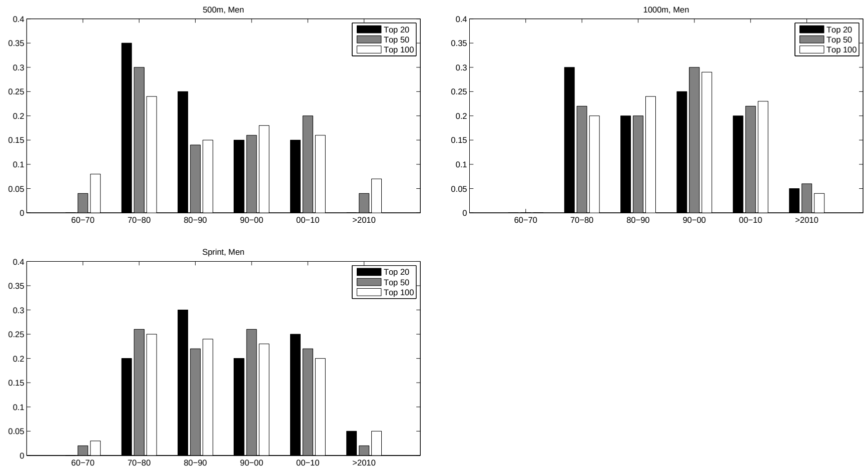



**Figure 2.18.** Distribution of top skaters over the seasons,

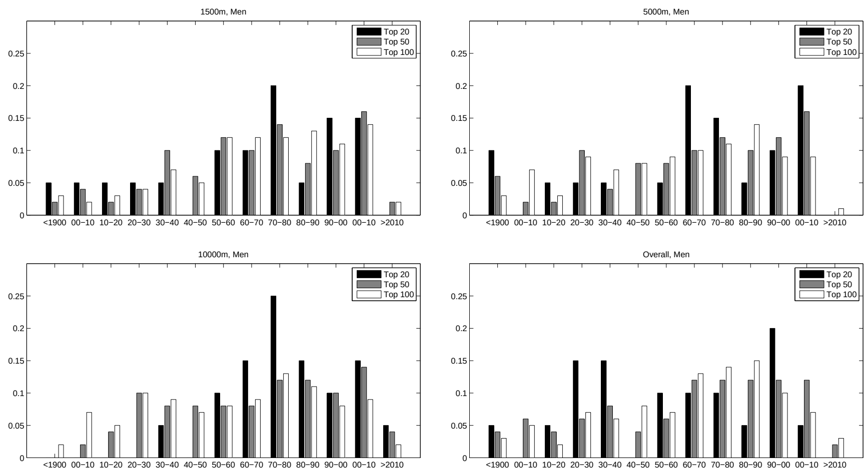



**Figure 2.19.** Distribution of top skaters over the seasons

Figure 2.18 shows that the top 20, top 50, and top 100 of the 1000m and the sprint discipline are uniformaly distributed. Of course, the last period, 2010-2011, is underrepresented, but all other four bars vary between 20% and 30%. For the 500m, the decade 1970-1980 is overrepresented, although one has to keep in mind that for the top 20 bars, a difference of 5% is equal to a difference of one skater. The figure shows that the period 1970-1980 delivers seven skaters to the top 20, and the period 2000-2010 only three skaters. This difference is also noticeable in the top 50 and top 100. We have two reasons for this phenomenon. First the introduction of the World

Sprint Championships in 1971, and secondly the short existence of the extra professional tournaments in 1973 and 1974 (see Section 2.2.2). Skaters can only be ranked in the 500m list if they participated in World Sprint Championships and Olympic Games, meaning that all 500m skaters who participated in the Olympic games of 1964 and 1968 also need a season from after 1971. The 1000m does not have this problem, since it was introduced on the Olympics in 1976. The introduction of the professional sprint tournament led to the fact that there were two world sprint tournaments in 1973 and 1974. On the other hand we can conclude that each of the four full decades have at least three skaters in the top 20.

The 1500m, 5000m, 10000m and the Overall histograms show more or less the same pattern. The bars of the first six decades (1890-1900 through 1950-1950) are all below the 10%, whereas the last six decades have all bars above the 10%. The main reason is that the participation level at the World Allround Championships and European Championships before 1950 was lower than during the decades after 1950 (see Section 2.3.5). Secondly, the period 1890-1950 faced two world wars in which no tournaments where organized. We know from Figure 2.4 that the participation level in 1890-1950 was almost half of that in 1951-2010, explaining the difference in number of skaters in the top 10, top 50 and top 100. However, with the exception of the 10000m, each decade still contains a black bar and is thereby present in the top 20 with at least one skater. This fact agrees with criteria RO 2.

The 10000m has no skater from before 1930 in the top 20. The 10000m is the 'most difficult' distance, especially in the early years when the weather circumstances could be very hard. Also, since in this period one could become allround champion by winning three distances, champions skipped the 10000m. The high peak in the period 1970-1980 is again due to the existence of the professional tournaments.

Since the bars of the top 50 and the top 100 in Figures 2.18 and 2.19 have, taking into account the number of active skaters per period, in all decades more or less the same height, we may conclude that each decade is represented in each list correctly.

**Sub-selection**

In this section we look at a subset from the final ranking and see if the ranking makes sense according to the best performances of the skaters on individual tournaments. We study the first twenty-two Dutch skaters that are ranked on the 1500m list within the top 100. Based on the results in their best three years we discuss the relative ranking positions. In Table 2.26 the ranking position, the ranking score, Olympic results, the allround tournament results, and the World Single Distance Championships results are presented.

Ard Schenk is the best Dutch 1500m skater. He won almost all 1500m in his best three years, and his score (0.37 better than the average top 5 result) is by far the best of all Dutch skaters. The Olympic champion of 1968, Kees Verkerk, is ranked below Ids Postma and Rintje Ritsma, who finished second and third respectively on

**Table 2.26.** Subranking of Dutch 1500m skaters

| Name | pos | score | OG | WChM | | | EChM | | | WSDCh | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Ard | 1 | -0.37 | ’72 | ’73 | ’72 | ’71 | ’73 | ’72 | ’71 | - | - |
| Schenk | | | 1 | 1 | 1 | 1 | 1 | 1 | 3 | - | - |
| Ids | 6 | -0.21 | ’98 | ’97 | ’98 | - | ’97 | ’99 | - | ’98 | ’99 |
| Postma | | | 2 | 1 | 1 | - | 1 | 3 | - | 2 | 1 |
| Mark | 10 | -0.16 | ’10 | ’04 | - | - | ’04 | ’05 | - | ’04 | ’05 |
| Tuitert | | | 1 | 2 | - | 1 | 2 | - | - | 2 | 2 |
| Rintje | 11 | -0.15 | ’98 | ’95 | ’97 | ’98 | ’95 | ’97 | ’98 | ’97 | ’98 |
| Ritsma | | | 3 | 1 | 5 | 2 | 1 | 2 | 1 | 1 | 4 |
| Kees | 12 | -0.14 | ’68 | ’66 | ’69 | ’68 | ’66 | ’69 | ’68 | - | - |
| Verkerk | | | 1 | 1 | 1 | 5 | 3 | 1 | 9 | - | - |
| Jaap | 13 | -0.13 | - | ”93 | ”95 | ”96 | - | - | - | - | - |
| Eden | | | - | 1 | 1 | 1 | - | - | - | - | - |
| Wim | 15 | -0.12 | ’52 | ’51 | ’53 | ’52 | ’51 | ’53 | ’52 | - | - |
| v/d Voort | | | 2 | 1 | 3 | 1 | 1 | 1 | 1 | - | - |
| Erben | 28 | -0.02 | ’06 | ’07 | - | - | - | - | - | ’03 | ’04 |
| Wennemars | | | 5 | 1 | - | - | - | - | - | 1 | 3 |
| Leo | 31 | -0.02 | ’92 | ’89 | ’91 | - | ’89 | ’91 | - | - | - |
| Visser | | | 3 | 1 | 5 | - | 7 | 1 | - | - | - |
| Falko | 37 | 0.02 | ’94 | ’93 | ’97 | - | ’93 | ’94 | ’97 | - | - |
| Zandstra | | | 3 | 2 | 15 | - | 1 | 2 | 3 | - | - |
| Jochem | 39 | 0.03 | ’02 | ’02 | ’05 | - | ’02 | ’05 | ’04 | - | - |
| Uytdehaage | | | 2 | 4 | 5 | - | 5 | 1 | 4 | - | - |
| Henk | 50 | 0.07 | - | ’62 | ’61 | ’60 | ’62 | ’61 | ’60 | - | - |
| v/d Grift | | | - | 2 | 1 | 5 | 2 | 3 | 12 | - | - |
| Frits | 65 | 0.12 | ’84 | ’82 | ’85 | ’84 | ’82 | ’85 | ’84 | - | - |
| Schalij | | | 10 | 2 | 10 | 2 | 1 | 2 | 2 | - | - |
| Hein | 69 | 0.16 | ’84 | ’85 | ’86 | - | ’85 | ’86 | - | - | - |
| Vergeer | | | 13 | 2 | 2 | - | 1 | 1 | - | - | - |
| Piet | 70 | 0.16 | ’10 | ’73 | ’76 | ’77 | ’73 | ’76 | ’77 | - | - |
| Kleine | | | 6 | 2 | 1 | 6 | 3 | 9 | 9 | - | - |
| Sven | 77 | 0.17 | ’10 | ’08 | ’09 | ’10 | ’08 | ’09 | ’10 | ’08 | ’09 |
| Kramer | | | 13 | 4 | 2 | 4 | 3 | 1 | 2 | 2 | 8 |
| Hilbert | 80 | 0.17 | ’84 | ’82 | ’84 | ’83 | ’82 | ’84 | ’83 | - | - |
| v/d Duim | | | 7 | 1 | 8 | 10 | 4 | 3 | 1 | - | - |
| Simon | 81 | 0.17 | ’06 | - | - | - | - | - | - | ’07 | ’08 |
| Kuipers | | | 4 | - | - | - | - | - | - | 5 | 8 |
| Rudi | 82 | 0.18 | ’64 | ’63 | ’66 | ’64 | ’63 | ’66 | ’64 | - | - |
| Liebrechts | | | 10 | 5 | 4 | 2 | 6 | 7 | 7 | - | - |
| Kees | 85 | 0.20 | ’52 | ’53 | ’52 | - | ’53 | ’52 | ’54 | - | - |
| Broekman | | | 5 | 4 | 5 | - | 2 | 8 | 7 | - | - |
| Martin | 87 | 0.20 | ’94 | ’96 | - | - | ’96 | - | - | ’96 | ’97 |
| Hersman | | | 8 | 7 | - | - | 3 | - | - | 3 | 4 |
| Harm | 90 | 0.22 | - | ’75 | ’74 | ’73 | ’75 | ’74 | ’73 | - | - |
| Kuipers | | | - | 5 | 5 | 5 | 2 | 5 | 5 | - | - |
| Jeroen | 93 | 0.23 | ’94 | - | - | - | - | - | - | ’95’ | ’97 |
| Straathof | | | 9 | - | - | - | - | - | - | 6 | 1 |

the Olympics. Verkerk has the disadvantage that in his Olympic year he was fifth on the World Championship Allround and ninth on the European Championships. Mark Tuitert, the Olympic champion of 2010, is ranked between Postma and Ritsma. Besides the Olympic victory, Tuitert only won the 1500m on the European Championships in 2004. On the other three races he became second.

The sixth Dutch man, Jaap Eden, won three 1500m’s during World Allround Championships and is ranked just above above Wim van der Voort. The tournament results of Postma, Tuitert, Ritsma, Verkerk, Eden, and Van der Voort do not differ much and also their final USS-ranking scores are almost the same. Ritsma and Postma have slightly better scores since they won their races with bigger leads.

Based on his Olympic result, the position of Erben Wennemars may look somewhat high. He participated only once in an allround tournament, but satisfies the tournament restriction (see Table 2.22) due to his participation in the World Single

Distance Championships. Wennemars won two of these three races and, next to that, he often finished within the first five in World Cup 1500m's.

Further down the ranking of Table 2.26, the Olympic performances become worse and less victories are observed at allround tournaments. Based on their Olympic result, Frits Schalij and Hein Vergeer are relatively high ranked. Their 10th and 13th places on the Olympics are worse than the Olympic results of skaters below them in this ranking. They owe the higher ranking position to their 1500m allround performances.

From Table 2.26 we may conclude that the ranking scores and the ranking positions of the Dutch skaters in the 1500m top 100 are as expected, if we look at the tournament results of these skaters in their best three years.

# 2.9 Scenario analysis

## 2.9.1 Tournament weights

In our USS-ranking model, we use CAV5-values as inputs for the performances of the skaters. The CAV5-values are derived from actual skating times. The influence on the final ranking score is determined by the value of the tournament weights. In this section we will show how the ranking scores are influenced by the tournament weight values. We only present and discuss the changes for the men's 1500m, 5000m, 10000m, and Overall disciplines.

For eight (ten) scenarios, the weight values of three tournaments, namely Olympic Games, World Allround Championships, and World Cup Competition, are changed; see Table 2.27

**Table 2.27.** Changing tournament weights

| **Scenario** | $L \in \{1500m, 5000m, 10000m\}$ | **Scenario** | $L = OV_M$ |
|---|---|---|---|
| 1 | - | 1 | $w_{\text{WACh,L}} = 10$ |
| 2 | $w_{\text{OG,L}} = 30$ | 2 | $w_{\text{WACh,L}} = 15$ |
| 3 | $w_{\text{OG,L}} = 50$ | 3 | - |
| 4 | $w_{\text{OG,L}} = 60$ | 4 | $w_{\text{WACh,L}} = 25$ |
| 5 | $w_{\text{OG,L}} = 100$ | 5 | $w_{\text{WACh,L}} = 30$ |
| 6 | $w_{\text{WACh,L}} = 15$ | 6 | $w_{\text{OG,L}} = 20$ |
| 7 | $w_{\text{WACh,L}} = 20$ | 7 | $w_{\text{OG,L}} = 30$ |
| 8 | $w_{\text{WC,L}} = 3$ | 8 | - |
| | | 9 | $w_{\text{OG,L}} = 50$ |
| | | 10 | $w_{\text{OG,L}} = 60$ |

changes in comparison to the weights given in Table 2.20

In order to understand how the weight changes will affect the various rankings, we have, for each scenario, calculated the share of a tournament weight in the total weight. In Table 2.28 the shares of each tournament in the total tournament weight is calculated for a skater who participated in one OG, three WACh and fifteen WCC races during his best three seasons. Note that we omit the weights of the ECh's and the WSDCh's. Table 2.28 shows that in Scenario 1, the original situation, the OG has about the largest impact on the score, namely 47%. In case of Scenario 2, the OG has the same influence as the average of the three results of the WACh in the three best seasons. For Scenario 6 and Scenario 7, the OG result has a lower share in the total weight than the three WACh results. For Scenario 8, the WC has the highest share. For the discipline Overall such a table is much harder to make, since the number of participations in the OG and the WACh varies much more between skaters on this discipline.

**Table 2.28.** The weight distribution

| Scenario | Tournament weight | | | | Share in total weight (%) | | |
|---|---|---|---|---|---|---|---|
| | OG | WACh | WCC | Total | OG | WACh | WCC |
| 1 | 40 | 10 | 1 | 85 | 47 | 35 | 18 |
| 2 | 30 | 10 | 1 | 75 | 40 | 40 | 20 |
| 3 | 50 | 10 | 1 | 95 | 53 | 32 | 16 |
| 4 | 60 | 10 | 1 | 105 | 57 | 29 | 14 |
| 5 | 100 | 10 | 1 | 145 | 69 | 21 | 10 |
| 6 | 40 | 15 | 1 | 100 | 40 | 45 | 15 |
| 7 | 40 | 20 | 1 | 115 | 35 | 52 | 13 |
| 8 | 40 | 10 | 3 | 115 | 35 | 26 | 39 |

Total= sum of tournaments weights of a skater skating one OG, three WACh and fifteen WCC races

The influence of the weights is analyzed by comparing, for all scenarios, the positions of the skaters in the original USS-ranking with the new positions as a result of the changed weights. In order to carry out this comparison, we will discuss the following situations: (1) The stability of the top 3, (2) the number of new names in the top 10, (3) the average number of position changes in the top 50, and (4) the distribution of the absolute position changes in the top 50.

For the stability of the top 3, we investigate whether or not the top 3 remains the same (-), change within (cw), or new names enter (ne). The average changes in the top 50 is measured by taking the mean over the absolute number of changes in positions. This mean varies between 0 (all skaters stay in the same position) and $x/2$, where $x$ is the number of skaters taken into account. So for the top 50, the mean can vary between 0 and 25. The distribution is depicted by spitting the top 50 into three groups, namely, skaters changing less than three positions (Position $< 3$), skaters changing between three to six positions (Position 3-6), and skaters changing at least six positions or more. The results for the 1500m, 5000m, and 10000m are summarized in Table 2.29, and will be discussed below.

**Changes in the top 3 and top 10** [7] Only when the weight of the WCC is changed a new skater enters the top 3 of both the 1500m and 5000m. For all other scenarios, the top 3 of these two distances remains the same. In case of the 10000m, Hjalmar Andersen takes over the first position of Johan Olav Koss when either the weight of the OG is decreased (Sc. 2), or the weight of the WACh is increased (Sc. 6 and Sc. 7).

[7] Our scenario analysis applies to the rankings of 2009, so that specific skater positions may slightly differ from the results in the appendix.

**Table 2.29.** Ranking position variability of 1500m, 5000m, and 10000m

| | Stability top 3 | | | | | | |
|---|---|---|---|---|---|---|---|
| | Sc. 2 | Sc. 3 | Sc. 4 | Sc. 5 | Sc. 6 | Sc. 7 | Sc. 8 |
| 1500m | - | - | - | - | - | - | ne |
| 5000m | - | - | - | - | - | - | ne |
| 10000m | cw | - | - | - | cw | cw | - |
| | - = no change, cw = position change within the top 3, ne = new name entering top 3. | | | | | | |
| | **Number of new names in top 10** | | | | | | |
| | Sc. 2 | Sc. 3 | Sc. 4 | Sc. 5 | Sc. 6 | Sc. 7 | Sc. 8 |
| 1500m | 0 | 1 | 1 | 2 | 0 | 1 | 2 |
| 5000m | 0 | 1 | 1 | 2 | 1 | 2 | 0 |
| 10000m | 0 | 1 | 1 | 2 | 0 | 0 | 0 |
| | | | | | | | |
| | **Average position change in top 50** | | | | | | |
| | Sc. 2 | Sc. 3 | Sc. 4 | Sc. 5 | Sc. 6 | Sc. 7 | Sc. 8 |
| 1500m | 1.6 | 1.2 | 1.9 | 4.1 | 2.8 (2.7) | 4.3 (3.9) | 2 |
| 5000m | 1.3 | 1.1 | 2 | 3.7 | 2.2 (1.6) | 4.3 (2.7) | 1.7 |
| 10000m | 1.8 | 1.2 | 2.3 | 4.3 | 2.5 (1.8) | 3.9 (2.9) | 0.4 |
| | (.) number of average changes when skaters with penalties are not taken into account. | | | | | | |
| | **Distribution position change top 50** | | | | | | |
| | Sc. 2 | Sc. 3 | Sc. 4 | Sc. 5 | Sc. 6 | Sc. 7 | Sc. 8 |
| Position <3, 1500m | 45 | 45 | 34 | 25 | 38 | 27 | 42 |
| Position 3-5, 1500m | 5 | 4 | 10 | 15 | 8 | 15 | 3 |
| Position >5, 1500m | 0 | 1 | 6 | 10 | 4 | 8 | 5 |
| | | | | | | | |
| Position <3, 5000m | 43 | 44 | 35 | 24 | 40 | 28 | 42 |
| Position 3-5, 5000m | 5 | 6 | 11 | 12 | 6 | 16 | 4 |
| Position >5, 5000m | 0 | 0 | 4 | 14 | 4 | 6 | 4 |
| | | | | | | | |
| Position <3, 10000m | 38 | 42 | 36 | 26 | 32 | 25 | 49 |
| Position 3-5, 10000m | 10 | 7 | 8 | 9 | 15 | 16 | 1 |
| Position >5, 10000m | 5 | 4 | 10 | 15 | 8 | 15 | 3 |

Koss has his best score on the OG, whereas Andersen scored better on the WACh. Table 2.29 shows that in all scenarios only at most two new names enter the top 10. The highest entry is made by Jochem Uytdehaage on the 10000m, when the OG-weight is increased to 100 (Sc. 5). He climbs from position 12 to position 6.

**Changes in the top 50** First we look at the average number of position changes. In case the weight of the OG is varied between 30 and 60 (Sc. 1, 2, 3, 4), the average change in position fluctuates between 1.1 and 2.3 for all three distances. However, when the weight is set to 100, the average position change increases to around four for all three disciplines. Table 2.28 shows that in case of Scenario 5, the OG weight represents almost 70% of the total weight, and therefore the total scores highly depend on the OG scores.

In case of Scenario 6, we also see an average of four position changes. Here the WACh is weighted twice as much as in the original model, so that the WACh weight has a higher impact on the final scores than the OG. In Scenario 6 and Scenario 7, we also observe the influence of the penalty score. Penalty scores partly depend on the WACh weight, and if the WACh weight increases so does the penalty score. This means that skaters with a penalty may drop too many places if the parameters of the penalty are not adjusted. For example, Julius Seyler drops 74 positions on the 5000m list in case of Scenario 7. Therefore, we also calculated the averages over the first 50 skaters without a penalty (values between brackets in Table 2.28).

The average values show that in case of Scenario 5 and Scenario 7, namely the scenario's in which the weight ratio between OG and WACh is changed the most, the rankings deviate the most from the original USS-rankings. This is also clear from the distribution of number of position changes of the top 50. Only for Scenario 5 and Scenario 7 more than ten skaters change more than five positions. In all other scenario's at least 34 of the 50 skaters remain within three positions of their original positions. Hence, if the hierarchical order of tournament weights remains unchanged, i.e., the OG is the most important tournament and is weighted slightly more than the three WACh's, no large position changes appear. However, when we make either the OG too important, or the WACh the most important, then the top 50 changes a lot. The top 3 and top 10 are quite robust to such large changes.

**Overall ranking** For the Overall ranking only changes in the OG and WACh weights are applied as these have the highest weights and most impact. In the first five scenarios, the WACh weight is increased from 10 to 30 with steps of 5 and in Scenario 6 through Scenario 10 the OG weight is increased from 20 to 60 with steps of 10. Recall that Scenario 3 and Scenario 8 are equal to the original situation. In Table 2.30 the results of the changes are presented.

For all scenario's the top 10 is represented by the same ten skaters as in the original ranking. Only in Scenario 5, one of the top 10 skaters (Gianni Romme) changes more than two positions, namely three. Heiden never leaves the first place. Ard Schenk only loses his second place to Johann Olav Koss (ranked third) if the WACh weight is deceased to 10 (Sc. 1), Koss loses is third place to the number four, Oscar Mathisen, if the OG weight is decreased to 20 (Sc. 5).

The top 50 is, as expected, more sensitive to changes in the WACh-weight than to changes in the OG weight. The WACh results have, in contrast with the single distance rankings, the largest share in the total weight, meaning that the ranking scores are more sensitive to changes in this weight. However, in case of both Scenario 1 and Scenario 5, still 72% and 82%, respectively, of the top 50 are ranked within five positions of their original position. For the top 100, we observe that this percentage decreases to 57% and 74%, repsectively.

The top 100 changes the most when the WACh weigth is set to 10 (Sc. 1). In this case, only 57 of the 100 skaters deviate at most five positions from their original position. This fact is, for a large part, caused by skaters with a penalty. For Scenario 1, skaters with a penalty replace original top 100 skaters, since their penalty is related to the WACh weight, which is now low. In the other situations skaters with penalty leave the top 100 due to an increasing penalty.

If either the WACh or the OG weight is increased, the impact on the ranking list

**Table 2.30.** Ranking position variability of Overall

| | Sc. 2 | Sc. 3 | Sc. 4 | Sc. 5 | Sc. 6 | Sc. 7 | Sc. 9 | Sc. 10 |
|---|---|---|---|---|---|---|---|---|
| | **Stability top 3** | | | | | | | |
| Top 3, OV men | cw | - | - | - | cw | - | - | - |
| | - = no change, cw = position change within the top 3, ne = new name entering top 3. | | | | | | | |
| | **New skaters in top 10** | | | | | | | |
| Top 10, OV men | - | - | - | - | - | - | - | - |
| | **Average change** | | | | | | | |
| OV men, Top 20 | 2.4 | 0.95 | 0.25 | 1.1 | 1.4 | 0.4 | 0.6 | 1.1 |
| OV men, Top 50 | 3.7 | 2.3 | 1.5 | 3.6 | 3.1 | 1.7 | 1.1 | 2 |
| | **Top 50** | | | | | | | |
| Position <3, OV men | 20 | 33 | 41 | 28 | 29 | 39 | 46 | 34 |
| Position 3-5, OV men | 16 | 13 | 6 | 13 | 11 | 7 | 3 | 14 |
| Position >5, OV men | 14 | 4 | 3 | 9 | 10 | 4 | 1 | 2 |
| | Top 100 | | | | | | | |
| Position <3, OV men | 31 | 57 | 68 | 45 | 44 | 66 | 78 | 45 |
| Position 3-5, OV men | 26 | 28 | 21 | 29 | 21 | 25 | 17 | 39 |
| Position >5, OV men | 43 | 15 | 11 | 26 | 35 | 9 | 5 | 16 |

OV = Overall, Sc. = Scenario

is less than when these weights are lowered with the same amount. Decreasing the WACh weight results in the fact that the other tournaments obtain more impact on the season scores, and therefore the set of best four seasons may changes in Scenario 1 for some skaters. For example, Yevgeni Grishin performed well at the ECh in 1961, but his lesser results on the WACh of the same season made the season 1961 his fifth best season. By lowering the WACh weight, his ECh score gains more impact on the season score, and makes it to one of his best seasons in Scenario 1.

We may conclude that if we decrease either the weight of WACh or OG by 50%, these tournaments loose their influence on the final ranking scores, and the Overall ranking list become significantly different from the original rankings. However, increasing the weights of these tournaments results in less changes.

To summarize, we may observe that in all four disciplines (1500m, 5000m, 10000m, and Overall), the top 3 and the top 10 remain more or less the same in all scenario's. The rankings are quite robust with respect to weight changes when the relative weight structure of the tournaments is kept the same. Scenario's that heavily change the relative tournament weight structure, show more volatile results, however still the majority of the names in the top 50 change less than three positions. The largest changes come from skaters with a penalty, because also their penalty value changes when the weight of the WACh is adjusted. The impact of the penalty function on the USS-rankings is explained in the following section.

### 2.9.2 Discussion of the penalty function

The restriction that skaters have to participate in at least one OG, can be 'overruled' by using the penalty function. The penalty function is used to rank skaters without an OG participation at the cost of a penalty that is added to their final score. In this section we justify how the parameter values of the penalty function (see Table 2.23) are chosen, we analyse how many skaters receive a penalty, and how the penalty has influenced the ranking positions.

**Justification of penalty parameter values** The penalty function is used for skaters without an OG and increases the final ranking score of this skaters such that they will be ranked lower. The penalty itself depends on the values of the parameters $\gamma_{cL}$ and $\beta_{cL}$, and on the total tournament weight value ($w_{kL}$) of the tournaments they have participated in in their best three (four) seasons; see Section 2.7.3. In Table 2.23, the values of parameters $\gamma_{cL}$ and $\beta_{cL}$ are given for each category $c$ and discipline $L$. Based the average difference in CAV5-values of two consecutive positions of the three important tournaments, we will discuss the impact of the penalty score.

Now, consider the results of all distance races of distance $d$. For each of these results the difference of any two consecutive (in the distance race result) CAV5-values is determined. The average of all these consecutive time differences are denoted by $ACD_d$. In Table 2.31 the $ACD_d$ values per tournament are presented.

**Table 2.31.** Average difference between consecutive positions (ACD)

| Distance | OG | WACh | WSCh |
|---|---|---|---|
| 500m | 0.03 | 0.13 | 0.04 |
| 1000m | 0.11 | - | 0.05 |
| 1500m | 0.11 | 0.13 | - |
| 5000m | 0.13 | 0.17 | - |
| 10000m | 0.21 | 0.18 | - |

Table 2.31 shows that, for example, during all 500m's of all OG's the ACD is 0.03 seconds. Furthermore, it shows that the differences are the smallest on the 500m, and the largest on the 10000m. The $ACD_{500m}$ value of the Olympics is almost equal to the $ACD_{500m}$ values of the WSCh. The$ACD_{500m}$ value of the WACh is higher, because in this tournament also non-sprinters participate. The WACh 500m results, however, are not taken into account in the 500m rankings.

The ACD values indicate how much a race-time has to be increased on average in order to lower a skater one position on average in that race ranking. Since the USS-ranking scores are averages of tournament scores the following happens.

Consider, for example, two 1500m skaters with exactly the same CAV5-values on their best three WACh's. Assume that skater A has skated an OG 1500m and that his CAV5-value on this OG is the same as the average CAV5-value of his three WACh's. So without a penalty both skaters would be ranked on the same position. If we would increasing the final average CAV5 score of the non-Olympic skater B with 0.11 (the $ACD_{1500m}$ of the OG), this would mean that his ranking position now drops, on average, 21 (namely the number of OG's) positions, because there are, on average 21 skaters with about the same score of skater A (see pre-condition 1, Section 2.6.5)).

In Table 2.23, it can be seen that the $\gamma_{cL}$-values are chosen about equal to the values of Table 2.31. This means that if a skater has missed an OG, but has participated in all other important tournaments, he will only receive the fixed penalty value, $\gamma_{cL}$, and increases his score such that he finishes, on average, one place lower on any distance race, so in total he drops about 21 positions.

In case a skater misses, besides an OG, other important tournaments, he receives one more time the value of $\gamma_{cL}$ per ten missing tournament weight points with respect to the required weight value, i.e., $\beta_{cL}$=$(0.1)\gamma_{cL}$. So for each ten missing tournament weight points, he drops one position extra on all his distance races. This means that if a skater misses, besides the OG, also one WACh result (WACh has a weight of 10), he receives twice the value of $\gamma_{cL}$ as penalty. Based on the values from Table 2.31, we see that his ranking score is set equal to a skater finishing, on average, two positions below him.

**Consequences of the penalty** Table 2.32 shows, for all disciplines, the total number of ranked skaters (row 1), and the total number of skaters with a penalty (row 2).

**Table 2.32.** Number of skaters with penalty

| | **500m** | **1000m** | **1500m** | **5000m** | **10000m** | **OV** | **SP** |
|---|---|---|---|---|---|---|---|
| # ranked | 356 | 266 | 550 | 506 | 287 | 377 | 238 |
| # with penalty | 116 | 88 | 223 | 204 | 126 | 128 | 97 |
| Percentage | 33% | 33% | 41% | 40% | 44% | 34% | 41% |
| # with penalty in: | | | | | | | |
| Top 50 | 3 | 5 | 4 | 3 | 4 | 0 | 2 |
| Top 100 | 7 | 10 | 11 | 15 | 24 | 3 | 6 |
| Top 200 | 22 | 40 | 49 | 57 | 71 | 23 | 61 |
| Percentage | | | | | | | |
| Top 50 | 3% | 6% | 2% | 1% | 3% | 0% | 2% |
| Top 100 | 6% | 11% | 5% | 7% | 19% | 2% | 6% |
| Top 200 | 19% | 45% | 22% | 28% | 56% | 18% | 63% |

#=total number of skaters

The table shows that quite a large number of skaters receive a penalty. The lower part of Table 2.32 shows that the majority of the penalized skaters are ranked outside the top 100, and for the larger ranking lists even outside the top 200. For the 1500m ranking, only four of the 88 (2%) penalized skaters are ranked within the top 50 skaters. In case of the Overall ranking all skaters with a penalty are outside the top 50.

The percentage of penalized skaters is slightly lower for the individual sprint distances compared to the 1500m, 5000m, and 10000m. The first reason for this fact is that the maximum number of participants on the OG 5000m and the OG 10000m is lower. For the 1500m the situation is somewhat special. This distance is during OG's accessible for all skaters, so also for skaters who normally only participate in 500m and 1000m races. These skaters have no other 1500m results and are not ranked, other 'real' allround skaters may loose their position to these skaters due to participation restrictions.

The other reason is that a lot of WACh skaters either choose to skate the 1500m or the 5000m/10000m. Although WACh skaters have results for all these three distances, they only have OG results for their most favorite distances, resulting in a penalty for the other OG distances.

Finally, we observe that a lot of penalized skaters, that are ranked below the 200th position, are skaters from the last two decades; they have participated mostly in the WCC and WSDCh. Especially for the 1500m and 5000m, we observe that, from the last 100 skaters, 75% receive a penalty.

**Rankings 'with' and 'without' the penalty** The first part of this section deals with the, above described, ACD-values. Based on these differences and pre-condition 1, we have indicated that the average decrease in positions for skaters, who only receive the fixed penalty value, is equal to the number of OG's, namely 21. Since most penalized skaters will also receive a variable penalty, based on the other tournaments they missed, the average decrease will be higher. In Table 2.32, the number of penalized skaters is listed. It can be observed that most of them are in the second part of the ranking lists. Below we will analyze how the penalty affects the ranking position, and show how many positions they actually drop.

The penalty influence is analyzed by comparing the positions of the penalized skaters in the USS-rankings with their positions in case they receive a penalty value equal to 0, i.e., $\gamma_{cL} = 0$ and $\beta_{cL} = 0$. In Table 2.33 the results are given. The first two rows show the average decrease in position of the first 100 penalized skaters and of all penalized skaters, respectively. The figures show that, for most disciplines, the average drop of the first 100 penalized skaters is higher than the average of all penalized skaters. Note that the 1000m and Sprint have no more than 100 penalized skaters. Many of the low ranked penalized skaters are already in the lower part of the USS-rankings, and thereby cannot drop much further. The values confirm more or less what is to be expected, namely a decrease of at least 21 positions. As said, since most penalized skaters receive besides the fixed penalty also the variable penalty, the average drop is more than 21 positions.

Only the 10000m shows a slightly lower value, but this can be explained by the fact that $\alpha_{I,100000m} = 0.1$, so only half the value of the ACD of the 10000m; see Table 2.31.

The second part of Table 2.33 shows the number of penalized skaters in the top 30 and the top 100 in case the penalty value is set to zero. The table shows that most of the penalized skaters leave the top 30, because the expected drop is at least 21 positions. For the top 100, we see that approximately 50% of the skaters with a penalty stay in the top 100.

**Table 2.33.** Comparing positions of penalized skaters with and without the penalty value

| | 500m | 1000m | 1500m | 5000m | 10000m | OV | SP |
|---|---|---|---|---|---|---|---|
| # skat. pen | 116 | 88 | 223 | 204 | 126 | 128 | 97 |
| Average drop first 100 pen. skat. | 28 | 23 | 34 | 25 | 17 | 30 | - |
| Average drop all pen. skat. | 30 | 23 | 26 | 15 | 14 | 25 | - |
| # pen. skat. in top-30 (no pen) | 3 | 5 | 5 | 3 | 2 | 3 | 2 |
| # pen. skat. in top-30 (with penalty) | 1 | 1 | 1 | 3 | 0 | 0 | 1 |
| # pen. skat. in top-100 (no pen) | 17 | 11 | 28 | 31 | 31 | 12 | 11 |
| # pen. skat. in top-100 (with penalty) | 7 | 10 | 11 | 15 | 24 | 3 | 6 |

(#= total number of, pen. skat. =penalized skaters)

### 2.9.3 Number of seasons

In the single distance USS-rankings the skaters are ranked based on three best seasons and the results on Olympic Games have a high weight, namely about 40 to 50 percent of the ranking score. The Sprint and the Overall USS-rankings are based on four years and in this rankings the world championships have the highest weight. Furthermore, one extra season decreases the score of a skater who peaked only during a short period of time. Due to the restriction that both the 500m and the 5000m have to be skated for the Overall discipline, specialist will also not dominate the Overall USS-ranking.

On the other hand, if we apply three instead of four seasons, it occurs that the difference are only small. In Table 2.34, the top 50 of the Overall USS-ranking based on four seasons is given. In the third column the ranking position based on three seasons is given, and in the fourth column the difference in positions. Column five and six give the USS-ranking scores obtained from four and three seasons, respectively.

Table 2.34 also shows that in the top 3, Sven Kramer loses his third place to Johann Olav Koss. Koss has one top season, namely the Olympic season of 1994 with three gold medals. Taking three seasons increases his score from -0.183 to -0.258. For Kramer the drop of one season has the opposite effect. His Olympic season is his worst of the four best seasons and therefore his score is worse if one of his better non-Olympic seasons is left out.

The top 10 of the Overall USS-ranking does not change, although Gianni Romme and Hjalmar Andersen are ranked higher under three seasons. Both are long distance specialist and benefit from the fact that one less season is used. The higher scores of specialist due to their Olympic score is exactly what we want to prevent with including another season.

In total, eight of the first fifty skaters change more than ten positions. Most of the top 50 skaters have more than three good seasons, and adding one more season has not much effect on the USS-ranking score. Among the skaters with only three good seasons is Peter Oslund, who climbs from position 25 to 11, when using three seasons. Ostlund's best four seasons have the scores -0.18, -0.16, -0.09, and 0.40. Note that his fourth score is much worse than the other three. Chad Hedrick, ranked on position 24, has a similar score profile.

The average change in absolute positions in the top 50 is 4.7, and 2.1 within the first twenty. about half of the skaters (24) change at most three positions and as said only eight skaters change more than ten. The extra season results in the fact that 66 skaters are not ranked anymore. Only, Vasili Ippolitvo (position 41) and Coen de Koning (position 91), are ranked within the first hundred if three seasons are taken. Only ten of the skaters that leave the ranking are within the first two-hundred.

The comparison for the overall ranking for women can be found in the Table A.15 of the appendix. The results are similar to the men's results. The number one, Gunda Niemann remains the best, the names in the top 10 do not change and the average drop in positions is equal to 2.9 in the top 50.

**Table 2.34.** Ranking 4 vs 3 seasons

| Skater | Pos. 4s | Pos. 3s | Dif. | Score 4s | Score 3s |
|---|---|---|---|---|---|
| Eric Heiden | 1 | 1 | 0 | -0.263 | -0.316 |
| Ard Schenk | 2 | 2 | 0 | -0.224 | -0.273 |
| Sven Kramer | 3 | 4 | -1 | -0.214 | -0.209 |
| Johann Olav Koss | 4 | 3 | 1 | -0.183 | -0.258 |
| Oscar Mathisen | 5 | 8 | -3 | -0.153 | -0.179 |
| Jaap Eden | 6 | 7 | -1 | -0.128 | -0.181 |
| Gianni Romme | 7 | 5 | 2 | -0.126 | -0.184 |
| Hjalmar Andersen | 8 | 6 | 2 | -0.122 | -0.181 |
| Clas Thunberg | 9 | 9 | 0 | -0.098 | -0.108 |
| Ivar Ballangrud | 10 | 10 | 0 | -0.097 | -0.104 |
| Rintje Ritsma | 11 | 12 | -1 | -0.044 | -0.057 |
| Ids Postma | 12 | 13 | -1 | -0.006 | -0.050 |
| Kees Verkerk | 13 | 17 | -4 | 0.007 | 0.004 |
| Oleg Goncharenko | 14 | 19 | -5 | 0.013 | 0.012 |
| Bernt Evensen | 15 | 18 | -3 | 0.026 | 0.004 |
| Tomas Gustafson | 16 | 15 | 1 | 0.030 | -0.020 |
| Birger Wasenius | 17 | 20 | -3 | 0.053 | 0.022 |
| Michael Staksrud | 18 | 27 | -9 | 0.058 | 0.080 |
| Knut Johannesen | 19 | 23 | -4 | 0.065 | 0.039 |
| Roald Larsen | 20 | 21 | -1 | 0.065 | 0.033 |
| Falko Zandstra | 21 | 22 | -1 | 0.066 | 0.036 |
| Sten Stensen | 22 | 16 | 6 | 0.076 | -0.013 |
| Shani Davis | 23 | 28 | -5 | 0.095 | 0.086 |
| Chad Hedrick | 24 | 14 | 10 | 0.111 | -0.043 |
| Peder Ostlund | 25 | 11 | 14 | 0.113 | -0.087 |
| Jochem Uytdehaage | 26 | 24 | 2 | 0.122 | 0.050 |
| Viktor Kosichkin | 27 | 30 | -3 | 0.151 | 0.106 |
| Oleg Bozhev | 28 | 32 | -4 | 0.156 | 0.122 |
| Yevgeni Grishin | 29 | 26 | 3 | 0.165 | 0.074 |
| Boris Shilkov | 30 | 39 | -9 | 0.167 | 0.154 |
| Gunnar Stromsten | 31 | 38 | -7 | 0.167 | 0.152 |
| Adne Sondral | 32 | 36 | -4 | 0.167 | 0.134 |
| Kay Arne Stenshjemmet | 33 | 42 | -9 | 0.170 | 0.161 |
| Rudolf Gundersen | 34 | 34 | 0 | 0.173 | 0.126 |
| Odd Lundberg | 35 | 40 | -5 | 0.177 | 0.160 |
| Leo Visser | 36 | 37 | -1 | 0.182 | 0.138 |
| Jan Egil Storholt | 37 | 47 | -10 | 0.185 | 0.193 |
| Fred Anton Maier | 38 | 25 | 13 | 0.200 | 0.063 |
| Enrico Fabris | 39 | 46 | -7 | 0.201 | 0.189 |
| Mark Tuitert | 40 | 35 | 5 | 0.201 | 0.131 |
| Nikolay Gulyayev | 41 | 29 | 12 | 0.210 | 0.101 |
| Hans Engnestangen | 42 | 57 | -15 | 0.213 | 0.220 |
| Martin Saeterhaug | 43 | 51 | -8 | 0.214 | 0.207 |
| Carl Verheijen | 44 | 44 | 0 | 0.216 | 0.176 |
| Charles Mathiesen | 45 | 59 | -14 | 0.225 | 0.221 |
| Franz Wathen | 46 | 33 | 13 | 0.228 | 0.125 |
| Piet Kleine | 47 | 45 | 2 | 0.234 | 0.180 |
| Hilbert van der Duim | 48 | 53 | -5 | 0.235 | 0.213 |
| Boris Stenin | 49 | 43 | 6 | 0.237 | 0.171 |
| Dag Fornaess | 50 | 54 | -4 | 0.242 | 0.214 |

Pos.= Position, s4= four seasons, s3= three seasons, Dif=Difference between Pos.s4 and Pos.s3

## 2.10 Conclusions

The idea of making an all time ranking list for speed skaters started when the president of the ISU, Ottavio Cinquanta, raised a question in an article in the Dutch speed skating journal Schaatssport Snoep (2004). "Is it possible to eliminate the influence of technological innovations that biases the currently used Adelskalender as all times ranking, and make a ranking that gives the 'old' champions the credits they deserve", he asked.

In this chapter, we described the problems that arise when comparing skating times from the period of 121 years of skating history. The race times have not only become better due to the improvement of the athlete's abilities but also due to the improvement in equipment, rinks, and trainings methods. In Section 2.3.1, we have discussed the impact of these changes on the skating times and world records. Part of the improvements can be related to Gould's hypothesis: the longer the sport is practiced under the same rules, the better the performances get, and the smaller the performance differences between top athletes become. The AV5-values, being the difference in race times with the average of the top 5 race times of that race, support Gould's hypothesis: the medium and variation of the AV5-value decrease over time. The actual testing of Gould's hypothesis with the AV5-values is biased by the fact that the number of tournaments and the participation level has changed over the years.

In Section 2.4 we formulated criteria for the performance score and ranking method in order to justify our ranking. In Section 2.5 we analyzed rankings based on victories or medal winning. We showed that the introduction of new tournaments, like the World Sprint Championships and the World Single Distances Championships, are a problem in these rankings. Since not all skaters were able to score points on these tournaments, they cannot be used.

One of the criteria was to eliminate the rink and equipment effects, and we used the AV5-values to do this, see Section 2.6.2. In the second step we have corrected the AV5-values for a continuous improvement effect, which we called the maturity effect. This effect covers the improvements Gould's hypothesis describes. The resulting values, called the CAV5-values, are used in our ranking model, called the Universal Speed Skating (USS) ranking.

In the USS rankings, the skaters are ranked on seven disciplines per gender, namely the five individual distances, the sprint discipline and the overall discipline. The final USS ranking score is based on the weighted average CAV5-values in their three (for the single distances) or four (for the sprint and overall discipline) best seasons, see Section 2.7.4. The weights are related to the tournaments, whereas the most important tournaments, such as the Olympics, have the highest weight. The three or four best seasons are chosen in such a way that the weighted CAV5-value is minimal under the restriction that, among others, at least one of the seasons contains a result from the Olympic Games.

The best male skater in the overall USS-ranking turns out to be Eric Heiden, and for the women this is Gunda Niemann. Most of the Olympic Champions are in the top of the USS-rankings. Furthermore, if we compare the top of the lists with rankings based on victories or medal winnings (see Section 2.5.3), the same skaters appear. In the top of the Overall USS-ranking each decade between 1890 and 2012 is represented by more or less the same number of skaters, supporting Cinquanta's idea that 'old' champions deserve a position in the top rankings.

### 2.10.1 Limitations and further research

In this final section of this chapter we will comment on the following four points.

- The early period;
- Tournaments versus seasons;
- Statistical testing;
- Gould's hypothesis

**The early period** A difficult problem was dealing with the low number of observations in the first thirty years (1892-1922) of speed skating . In this period only two tournaments where organized, the participation level was low, and the skating times fluctuated heavily. Furthermore, since traveling was much harder those days, not all top skaters could be present always, and so many participants came from the organizing country. For example, eleven of the sixteen skaters of the first World Allround Championships in Amsterdam were Dutch.

No wonder, that our scenario analyses (see Section 2.9) we found that the ranking positions of skaters from this period are the least robust against parameter changes. Actually, many race times from this period are outliers. Hence, for more robust results it is better to start the database in 1924, the first OG season.

**Tournament vs Season** The three/four best seasons are chosen in such a way that the average weighted score of the CAV-values in those seasons is minimal under the restriction that the skater has participated in a number of compulsory tournaments. All results from a chosen season are included, inclusively the worst ones, resulting it the fact that a good season can be left out of the best three/four seasons due to one race result. This problem can be avoided by taking the individual tournaments as decision variable instead of complete seasons. We then still would require at least one OG, plus two WACh or two WSDCh races, but other tournaments are free to choose. The drawback of this choice is that in case the OG score plus the two WACh scores are the best scores of a skater, then the addition other tournament scores would worsen the final score and are therefore not taken into account in the ranking procedure. This means that we would have to include more restrictions regarding either the other tournaments, or the number of races used for the calculation.

**Statistical testing** We already applied sensitivity analysis on a number of input parameters, such as the tournament weights and the penalty parameters. For further research it would be an interesting direction to test if the differences between ranking scores of the skaters are significant. In the sensitivity analysis on the penalty parameters we already have seen that the differences between the lower ranked skaters are small and small changes influence these rankings. On the other hand the sensitivity analysis on the tournament parameters showed us that the top ranked skaters are rather robust.

**Gould's hypothesis** A large part of the correction of absolute skating times is based on the hypothesis of Gould. Many factors that have improved the performances continually over time, like training methods, diets and financial resources, are labeled as the maturity factors. We have assume that the maturity rate has the same influence on all skating times and is constant within one season. The skating population and participation level within tournaments, however, has changed over the years, and interfere with the effects described by Gould's hypothesis. Therefore, our analysis in Section 2.6.3 is just an indication of the existence of Gould's hypothesis.

Further research regarding the hypothesis should give more insight in all effects that play a role and accept or the reject the hypothesis based on formal tests. A more quantitative estimation of the exact influence of rink and equipment innovations, such as the introduction of the klapskate, would also be of interest. We have taken in our analysis the differences between times within tournament races and neglected the exact effects. In Kamst (2010), for the period 1994 through 2009, estimations are made to identify the influence of rink, season and month effects on skating times. The resulting corrected times are used to compare young talented skaters and to identify talents. It is found, for example, that skating times on the high altitude rink of Calgary are on average 2.7% faster than on the sea level rink Heerenveen. These kind of estimations can also help in further investigations of Gould's hypothesis

Chapter 3

# Performance Comparisons of Dutch Professional Soccer Players

## 3.1 Introduction

This chapter deals with the problem of identifying, measuring and comparing the performance of individual team members. A major question concerns the measurement of the individual contribution to the team performance. We use an all time ranking methodology in which performances of individual team players are compared. The actual performance of the individual players is derived from the team performance. We restrict ourselves to Dutch professional soccer players. The analysis will lead to an all-time ranking of these players. Our data set starts with the season 1955/56, since this season marks the start of professional soccer in the Netherlands. The ranking is mainly based on the actual playing time of the players during the important international matches and tournaments. The main reason of this choice is that actual playing time reflects the expertise and judgments of coaches, who can be seen as the best judges of players performances. Points are awarded for winning matches and advancing in tournaments, taking into account the importance of the tournaments, the number of participants and the stage of the tournament. Furthermore, we will make corrections on playing times. For example, goalkeepers tend to be less often substituted and have fewer injuries in comparison to forwards.

Maybe of all sports, soccer knows the largest number of types of rankings. There are both individual and team rankings. Examples of individual rankings are top scorers rankings and player-of-the-year rankings. Team rankings are used during tournaments and competitions. Most of these rankings are made for one season. There are fewer rankings that cover larger time periods, mainly because the comparison is more complicated due to changing circumstances, tournaments and rules, and the fact that players from different decades did not play against each other. There are a few all time team rankings but for individual players no such ranking exist. In this chapter a model is presented that compares individual soccer players from the period 1956-2011 and ranks them based on their performances in international club an country matches.

## 3.2 Competitions and tournaments

There is a wide range of tournaments and competitions, while there are mainly two types of matches, namely matches between clubs and matches between countries.

### 3.2.1 Club team football

Professional soccer teams participate in a national competition, such as the Dutch Eredivisie, the English Premier League, or the Spanish Primera División. Additionally, the best clubs of the European countries play for European cups, such as the UEFA Champions League and the UEFA Europa League. In other continents similar tournaments are organized, for instance the Copa Libertadores in Latin-America.

The number and type of European club tournaments has not always been the same. Table 3.1 shows an historical overview of the most important tournaments. This overview starts in the season 1955/56 (seasons are denoted by its second half year, i.e. season 1955/1956 is denoted as 1956). Below we present a more detailed description of the tournaments listed in Table 3.1.

**Table 3.1.** European club tournaments

| Tournament | Season |
|---|---|
| European Cup | 1956 - 1992 |
| Champions League | 1993 - current |
| Inter-cities Fairs Cup | 1956 - 1971 |
| UEFA Cup (UEFA Europa League) | 1972 - current |
| UEFA Cup Winners' Cup | 1961 - 1999 |
| UEFA Super Cup | 1972 - current |

- **European Cup and Champions League (EC1)** The European Cup (in Dutch: *Europacup I*) was founded in 1955. Initially, only country champions could participate, and it was organized as a knock-out competition. In 1992 the European Cup was renamed UEFA Champions League and also non country champion teams could enter the tournament. One season earlier, the knock-out system was replaced by a system with both group stages and knock-out stages. The format was changed many times since then. This tournament has always been considered as the most prestigious club tournament in Europe.

- **UEFA Cup Winners' Cup (EC2)** The EC2 (in Dutch: *Europacup II* or *Beker voor Bekerwinnaars*) was founded in 1960 and open to country cup winners. The tournament was considered the second most important, but its status declined in the nineties, when the Champions League attracted most of the money and the media attention. During the season 1998/99 the tournament was organized for the last time.

- **Inter-Cities Fairs Cup (IFC)** The IFC (in Dutch: *Jaarbeursstedenbeker*) has been organized between 1955 and 1971. It was open to teams from cities that hosted trade fairs. Initially the ranking of a team in its national league was not taken into account. During the last three years of its existence, teams could qualify only based on their national competition ranking.

- **UEFA Cup (EC3)** In 1971 the IFC was replaced by the EC3, which was organized as a knock-out competition. Teams qualified based on results in the national leagues and cups. Initially, the EC3 was considered to be the third most important European tournament, after the EC1 and the EC2. When the latter was disbanded in 1999, the EC3 became the second most important one. Since 2004/05 the tournament starts with a group stage, followed by a knock-out competition. Currently it is called the UEFA Europa League.

- **UEFA Super Cup (SC)** The SC is a yearly match between the winners of the two most important European cups. From 1972 through 1999 the match was played between the winner of the EC1 and the winner of the EC2. After 1999 the latter was replaced by the winner of the EC3. Until 1997 the SC was organized over two matches; since then there is only one match.

### 3.2.2 National team football

A national team of a country is considered to be a selection of players that forms the best team of that country. For European teams there are two important international tournaments, namely the World Cup organized by the FIFA, and the European Championships organized by the UEFA. Both tournaments are organized once in four years with two years in between. For the other continents there are similar tournaments to the European Championships.

- **FIFA World Cup (FWC)** The FWC is the most important tournament between countries and is organized every four years since 1930 with the exception of 1942 and 1946. The tournament typically consists of a group stage and a few knock-out stages.

- **UEFA European Championship (UEC)** The UEC started in 1960, and is organized in a similar way as the FWC. The UEC can be seen as the continental equivalent of the FWC.

Besides these two important tournaments there are qualifier matches for the FWCs and the UECs, and friendly matches.

### 3.2.3 Team rankings

In tournaments or competitions teams play against each other, and at the end of the competition the final ranking reflects the overall performances of the teams. For each individual tournament and competition, there are rules used to quantify game results and to make team rankings. Usually winning a match results in three points, losing no points, while a draw results in one points. Before 1994, however, winning the match resulted in only two points. For teams, there are a few rankings that compare results of different competitions and tournaments. We mention the most important ones. For the national team we have the

- **FIFA World Ranking** The FIFA World Ranking was introduced in 1993 and reflects the mutual strength of national football teams. For each match played a country receives points, where the number of points depends on four factors: the result of the match (won, lost, or draw), the importance of the match, the ranking position of the opponent, and the confederation to which the opponent belongs. The exact rules can be found on the website of FIFA (see, FIFA (2011)). Despite some skepticism in the beginning, the FIFA World Ranking is widely considered to be the most important indicator for the strength of national football teams.

- **UEFA National Team Coefficient** The UEFA also has a ranking of national teams based on the UEFA National Team Coefficient. This coefficient is calculated using similar rules as the FIFA World Ranking. These rules can be found on the website of UEFA (see, UEFA (2011)). Since the UEFA ranking consists of only European teams it is considered less important than the FIFA ranking. The main purpose is to seed the calculations for the qualification for UECs.

**Table 3.2.** National team rankings, Nov 2011

| | **FIFA World Ranking** | | **UEFA Nat. Team Coeff.** | |
|---|---|---|---|---|
| Rank | Team | Points | Team | Points |
| 1 | Spain | 1624 | Spain | 39,964 |
| 2 | Netherlands | 1425 | Netherlands | 38,294 |
| 3 | Germany | 1352 | Germany | 37,821 |
| 4 | Uruguay | 1230 | Italy | 35,838 |
| 5 | Brazil | 1144 | England | 34,819 |

For club teams there exist a number of custom made club rankings, mostly used on gambling sites. One of the most sophisticated rankings is developed by Infostrada (see, Infostrada (2011)), named the Euro Club Index, which estimates the club strength based on results from July 2007 until now in national leagues, cup matches and in the international UEFA tournaments.

On the Rec. Sport Soccer Statistics Foundation (RSSSF) website (see, RSSSF (2011)) a list is presented in which all club teams are ranked by the number of points they gained in the international UEFA tournaments since 1956. The system, developed by Jerome Faugeras (see, Faugeras (2011)), gives each club two points for winning a match and one point for a draw. Moreover, reaching the quarter finals of any cup (either knock-out round or group stage of eight clubs) yields one bonus point. This also holds for reaching the semi-finals, reaching the final, and winning the competition. In Table 3.3 the top five of both lists is presented.

**Table 3.3.** Club team rankings, Nov. 2011

| | **FIFA World Ranking** | | **UEFA Nat. Team Coef.** | | |
|---|---|---|---|---|---|
| Rank | Team | Points | Team | Matches | Points |
| 1 | FC Barcelona | 4271 | FC Barcelona | 458 | 703 |
| 2 | Real Madrid | 4114 | Real Madrid | 434 | 659 |
| 3 | Manchester United | 3939 | Juventus | 386 | 579 |
| 4 | Chelsea | 3520 | Bayern München | 373 | 566 |
| 5 | Bayern München | 3516 | AC Milan | 336 | 497 |

## 3.3 Ranking model

This chapter contains a mathematical model for ranking soccer players. We start by formulating a set of pre-conditions, being a set of criteria that need to be satisfied by the final ranking.

### 3.3.1 Ranking criteria

The following pre-conditions are used to check the quality of the final all-time ranking.

***Pre-condition 1: "Performance".*** *Players from the same time period should be ranked according to the ranking in that period.*

Athletes that have played against each other are, of course, easier to compare than athletes from different time periods, that never have played against each other. Also competing circumstances may have changed. Although this criterion is rather trivial, it does not mean that it is easy to compare performances in a fixed time period.

***Pre-condition 2: "Field position".*** *In the top of the ranking the distribution of field positions should be more or less according to the ratio of the most common playing formation (4-3-3 or 4-4-2.)*

Every team has one goalkeeper and ten other field players. Depending on the team strategy, there are for instance four defenders, four midfielders, and two forwards (4-4-2), or four defenders, three midfielders and three forwards (4-3-3). Therefore we may expect that in the top of the ranking each position is represented according to this distribution and not that one position dominates another.

***Pre-condition 3: "Tournaments"*** *The ranking does not depend on whether or not tournaments are organized in a season.*

Tournaments come and go over the seasons; see Table 3.1. When a tournament is not organized anymore in a season, then players active in that season should not have a disadvantage of not having the opportunity to be active in that tournament.

***Pre-condition 4: "Tournament structure".*** *The ranking does not depend on changing tournament structures over the seasons.*

Tournament structures differ a lot over time. Usually, there are knock-out stages and group stages. The structure and the number of participants have been adjusted many times for almost all tournaments. We assume that the importance of obtaining a certain result in a tournament does not depend on the format of the tournament.

Winning the UEFA Cup is equally valuable in 2004 as it is in 2005, although in 2004 the tournament was organized as a knock-out competition, while in 2005 a group stage was added.

### 3.3.2 Ranking model

One of the most difficult aspects of comparing performances and results of soccer players is the fact that soccer is a team sport. Winning a match or a tournament is always the result of the team as a whole. How can we measure the contribution of an individual player to the team result? Of course, scoring a goal and giving an assist are positive contributions. On the other hand, red or yellow cards can be seen as negative valued indicators. Being in the right position to score or to give an assist are examples of performance contributions that highly depend on the other players. In order to make a fair comparison we need to use much more information about passes, tackles, and all other player's actions. However, collecting data on all these attributes needs a highly complicated model and is actually almost impossible to build. Moreover, this kind of data is not available for all players.

The above discussion leads us to the basic idea of using team results, rather than specific individual performance results. Basically, we assume that each player in the team contributes equally to the final result of the team. A drawback could be that for one single match, it may happen that a player obtains not enough or too much appreciation, at least different from what he actually deserves. However, on the scale of a complete career these deviations are assumed to cancel out. So our ranking method will be based on team results, rather than on the individual technical skills displayed during the match. This choice also means that we consider the selection policy of the coach as a major gauge for determining the performance level showed and realized by the player.

**Data and notations**

In order to determine the number of points players may receive for the final result of a match, we first have to decide which matches and tournaments are taken into account. We have decided to use only performances on the 'international level'. National competitions are seen as selection rounds for the international tournaments. A major drawback of this choice is that only players that have acted in such competitions will appear in our all time ranking. This implies that we restrict ourselves to the top seeded Dutch soccer players, which have played at least one official international match in the national team. The international competitions that are used are listed below.

- FIFA World Cup (FWC);
- UEFA European Championship (UEC);
- Qualifier for FIFA World Cups and UEFA European Championships (Q);

- European Cup I and UEFA Champions League (EC1);
- UEFA Cup Winners' Cup (EC2);
- Inter-cities Fairs Cup (IFC);
- UEFA Cup and UEFA Europa League (EC3);
- UEFA Super Cup (SC).

As is mentioned already, we start with the season 1955/1956, the first season of Dutch professional soccer. The last season for which we have collected data is 2010/2011. Recall that we refer to the seasons as "years"; so for instance the year 1956 corresponds to season 1955/1956. For each player in our data set, we collect data on the following dimensions.

- Name player;
- Player's field position;
- Date match;
- Names both teams in match;
- Name tournament;
- Stage in tournament;
- Match result;
- Minutes played by player.

We define the following index sets.

| | | |
|---|---|---|
| $Y$ | = | $\{1956, \ldots, 2011\}$, the set of seasons; |
| $P$ | = | set of players; |
| $T$ | = | $\{FWC, UEC, Q, EC1, EC2, IFC, EC3, SC\}$, set of tournaments; |
| $\mathcal{I}_{ty}$ | = | $\{1, \ldots, I_{ty}\}$, set of stages in tournament $t$ in season $y$; |
| $\mathcal{J}_{tyi}$ | = | $\{1, \ldots, J_{tyi}\}$, set of matches in stage $i$ in tournament $t$ in season $y$; |
| $m_{tyijp}$ | = | number of minutes player $p$ has played in match $j$ in round $i$ of tournament $t$ in season $y$. |

Finally, we define the set of field positions as

$$F \in \{\text{GK (Goalkeeper)}, DF \text{ (Defender)}, MF \text{ (Midfielder)}, FW \text{ (Forward)} \} .$$

### Performance score

In this section, we introduce the concept of Performance Score: It is a total number of points any soccer player can collect during his career. The actual calculation of this concept will be explained hereafter. Performance Scores are based on the following assumptions.

***Assumption 1: "Team"*** *The Performance Score of a player in a match is determined by specific characteristics of this match.*

The basic idea is that players receive points for each match they play. This number of points will depend on a number of match specific factors which we will describe below.

**Playing time** In principle, all players of the team receive the same number of points for each match they play. However, we correct for the number of minutes a player actually plays in the match. The fact that we take into account the number of playing minutes is expressed in the following assumption.

***Assumption 2: "Playing time"*** *The contribution to the Performance Score of a player in a match is proportional to his playing time in that match.*

Let $t \in T$, $y \in Y$, $i \in \mathcal{I}_{ty}$, $j \in \mathcal{J}_{tyi}$, and $p \in P$. Then we define

$$\begin{aligned} \frac{m_{tyijp}}{90} &= \text{the contribution of player } p \text{ in match } j \\ &\quad \text{of stage } i \text{ in tournament } t \text{ in season } y. \end{aligned} \tag{3.1}$$

In general, a soccer match consists of 90 minutes. During knock-out matches, however, extra time can be added when during the 90 minutes the match is in a draw and a winner is needed. If player $p$ has played more than 90 minutes, we set $m_{tyijp} = 90$, so that the playing proportion always stays between 0 and 1. If player $p$ played less than 90 minutes (this also holds for players that where substituted during overtime) $m_{tyijp}$ is simply the number of minutes played.

**Match points** The Performance Score will also depend on the team result, i.e., if a match is either won, lost, or ended in a draw. Nowadays, in soccer competitions a team receives three points for a victory, one point for a draw, and zero points for a lost match. Before 1994, a won match corresponded to two points instead of three. In our ranking model we deviate from these valuations, because of the fact that a player who was in the line-up of a lost match is considered 'better' than a teammate, who did not receive playing minutes from the coach. When we would assign the value of 0 for each lost match, the Dutch players of the World Cup Final of 1974 would not obtain credits for playing this prestigious final. Therefore, we add one point to the usual point distribution, where we use the following variant.

***Assumption 3: "Match points"*** *The Performance Score of a player depends on match results.*
The multiplier for the match results is defined as follow: The number of playing minutes is multiplied by $r$, defined by:

$$r = \begin{cases} 1 & \text{if the match is lost;} \\ 2 & \text{if the match has ended in a draw;} \\ 4 & \text{if the match is won;} \\ 0 & \text{otherwise.} \end{cases}$$

The above rules do not take into account the situation when the match is finished with penalty shoot outs. In case the match was ended in a draw during a non-final stage, we assign two points to both teams. The team that has won the shootout proceeds to the next round and will receive there at least one more point. In case of a final during a tournament, the winner always receives three points and the loser one point, even if the final result is achieved after a shootout.

**Tournament importancy** A difficult problem concerns the rating of tournaments. How much more important is the EC1 than, for example, the EC3? And is a UEC between countries more important than winning the EC1 between clubs? As can be seen from Table 3.1, the distribution of European club tournaments over the years has not always been the same. According to Pre-condition 3 we have to include corrections in our model for these differences. For instance, players who were not active between 1961 and 1999 should not have a disadvantage of the fact that the EC2 was organized between 1961 and 1999. Therefore, we will define the European club tournament parameters in such a way that in each period more or less the same total number of European club tournament points can be assigned.

Unlike the European club tournaments, that are organized each year, both the FWC and the UEC are organized each four years, giving players the opportunity to participate in these tournaments at least once during their career. Moreover, we assume that the relative importance of the various tournaments does not change over time. For each $t \in T$ and $y \in Y$, we define

$$V_{ty} \quad = \quad \begin{cases} \text{the mutual relative importance} & \\ \text{of tournament } t \text{ in season } y & \text{if } t \text{ is organized in } y \\ 0 & \text{otherwise.} \end{cases} \qquad (3.2)$$

So, in case a tournament $t$ is not organized in season $y$, we define $V_{ty} = 0$. But how to determine the other values for the parameter $V_{ty}$? In the FIFA World Ranking for national teams (see Section 3.2.3), the FIFA uses parameters that reflect the importance of matches. Table 3.4 shows the FIFA values together with our values for the relative importance parameters. Our values of $V_{ty}$ deviate from the FIFA values. Friendly Matches (FM) and Qualifiers (Q) both obtain lower values, while the FWC and the UEC obtain much higher values. The reason is that we are interested in the matches where players have to perform at their top level and we assume that this is the case during FWCs and UECs and not during FM and Q matches. In Section 3.5,

**Table 3.4.** Relative importance interland matches/tournaments

| Type of match (t) | FIFA | $V_{ty}$ |
|---|---|---|
| FM | 1.0 | 0 |
| Q | 2.5 | 0.5 |
| UEC | 3.0 | 7 |
| FWC | 4.0 | 10 |

we will apply sensitivity analysis on the relative importance factors from Table 3.4.

For the European club tournaments such weight factors are not officially determined. There is no discussion about the EC1 being the most prestigious championship for club teams in Europe. However, for the EC3 and the EC2 the opinions differ. One may expect that the EC2 is considered to be more important than the EC3. However, based on the strength and the number of participating teams one may claim that the EC3 is more important. The EC2 was a tournament between cup winners, whereas the EC3 is a tournament between teams that finished in the top of their national competition. In several occasions, the team that won the cup was ranked lower in the national competition than a team that participated in the EC3. We assume that both tournaments are equally important, also because after 1999 they are merged.

To express the superiority of the EC1, we multiply matches from this competition by two times the value of the matches from the EC2 and the EC3. Since the IFC is a predecessor of the EC3 we consider both tournaments of equal value. The SC is considered quarterly as important as both the EC2, the EC3, and the IFC. This is because the SC consists of only one match and is a bit of an extra prize for teams who already won a tournament. The importance factors for all European club tournaments are listed in Table 3.5.

**Table 3.5.** Relative importance European club tournaments

| Tournament (t) | $V_{ty}$ |
|---|---|
| EC3 and IFC | 2 |
| EC2 | 2 |
| EC1 | 6 |
| SC | $\frac{1}{2}$ |

In our ranking model we will have two options for these relative importance factors, namely (1) each tournament has a fixed weight for the entire period, or (2) each season has the same total number of points for all European club tournaments, and the tournament points are adjusted according to the number of tournaments. In case of option (1), we have the following assumption.

***Assumption 4a: "Fixed tournament weights"*** *The weight of each European club tournament is the same for all seasons.*

We define for each $t \in T$ and $y \in Y$,

$$\tau_{ty} = \begin{cases} V_{ty} & \text{if } t \text{ is organized in } y \\ 0 & \text{else} \end{cases} \tag{3.3}$$

An argument against this option is that the most prestigious tournament, the $EC1$ (currently called Champions League) has become more competitive due to the increasing number of participating teams. Nowadays, the champions of more countries join this tournament, while also the number two and three of the best performing competitions are given the opportunity to enter the Champions League. Due to the increase of teams, the number of rounds also has increased. There are even qualifying rounds before the actual tournament starts. So we might argue that the Champions league deserves a higher weight after the increasing number of participating teams.

On the other hand the opportunity of entering the highest international club tournament has also increased. Before the introduction of the Champion League a team had to become champion of its national competition whereas today the number two and even sometimes the number three still have a chance of entering the tournament and win it. This means that the current best players almost every year play in the Champions League.

In the second option we take in to account that the number of club tournaments has changed over the seasons. As the number of tournaments changes, the possibility to win a tournament changes. In order to compensate for this, we introduce a parameter that takes care of the fact that each season has the same weight concerning the European competitions, meaning that in each season the total number of points that can be earned from European clubs is a constant.

***Assumption 4b: "Season balance European Cups"*** *The total weight of all European club tournaments is the same for all seasons.*

Based on this assumption, we define,

$$c = \text{the total number of points that can be earned in one season from all European tournaments in that season.} \tag{3.4}$$

For each $t \in T$ and $y \in Y$, we define

$$\tau_{ty} = \begin{cases} c\frac{V_{ty}}{\sum_{t \in T \setminus \{WC,EC,Q\}} V_{ty}} & \text{if } t \in T \setminus \{WC, EC, Q\}\,; \\ V_{ty} & \text{if } t \in \{WC, EC, Q\}\,. \end{cases} \tag{3.5}$$

The above definition means that, for each season $y$, $\tau_{ty}$ is the fraction of the weight of European cup tournament $t$ in that season and the sum of the weights for all European cups in that season. For the international tournaments, UEC, FWC, and Q, we do not need to correct for the seasons, so that for each season $y$ with $t$ = UEC, FWC, Q, $\tau_{ty}$ is equal to the relative tournament weight $V_{ty}$. The use of equation (3.5) will be illustrated later on in Example 1.

To complete the weighting system, we have to determine a value for $c$. In order to do so, we need to know the relative importance of the European club competitions and international tournaments. The first differences between the tournaments is that the international tournaments are only organized once in the four seasons, whereas the European club competitions are organized each year.

In terms of numbers of spectators, international matches score much higher than European club matches. In 2008, according to the *Stichting Kijk Onderzoek* ( see Kijkonderzoek (2011)) six out of the ten best viewed television programs in the Netherlands were matches of UEC 2008 (one match with 7,572,000 viewers), while the first EC1 match in this ranking was in position 57 (with 2,213,000 viewers). In 2010 eight of the ten best viewed programs were matches of the FWC 2010, whereas the final match between the Netherlands and Spain had 8,513000 viewers. The final of the EC1 in that season was ranked on the 64 position with 2,496,000 viewers. Although the exact number of viewers can fluctuate from season to season and highly depends on the success of the own club teams, the difference in viewers gives an indication in the difference of importance of the tournaments.

Based on frequency and the viewers attention, we choose $c = 10$, meaning that the total of European club tournaments in one season is equally important as the FWC. The tournament parameter values resulting from this value are listed in Table 3.6. All values are calculated in the same way as is done in Example 1. Since there are several club tournaments in one season, the importance factor of any individual club tournament is less than that of a FWC and in the period 1961-2000 even less than that of an UEC (see Table 3.6).

**Example 1.** *During the season 2007/08 the following tournaments are organized: EC1, EC3, SC, and UEC. In case of the UEC, $\tau_{ty} = V_{ty}$ for each $y$, so that $\tau_{EC-2008} = V_{EC-2008} = 7$ (based on Table 3.4). For the other tournaments, the mutual relative importance factors can be found in Table 3.5. Then, $V_{EC1-2008} = 6$, $V_{EC3-2008} = 2$, and $V_{SC-2008} = \frac{1}{2}$, so that $\sum_{t(2008)} V_{ty} = 8\frac{1}{2}$. Using (3.5) with $c = 10$, we find that $\tau_{EC1-2008} = 7\frac{1}{17}$, $\tau_{EC3-2008} = 2\frac{6}{7}$, and $\tau_{SC-2008} = \frac{10}{17}$.*

It can be seen that, with the exception of the period 1956-1960, the current EC1 is relatively more important than it was before. Since the number of club tournaments decreased from three to two, the possibility of winning a tournament has decreased. Winning a tournament has become more difficult and therefore is rewarded slightly more in the equal season weight option. In the fixed weight option no distinction is made and each EC1, EC2 or EC3 victory is seen as equally impressive.

In Section 3.5.2 we compare the result from using either assumption 4a or 4b, i.e. either we use equation 3.3 or 3.5.

**Table 3.6.** Tournament weights ($\tau_{ty}$)

| | EC3 | IFC | EC2 | EC1 | SC | WC | EC | Q |
|---|---|---|---|---|---|---|---|---|
| **Period** | Season budget weights | | | | | | | |
| **1956 - 1960** | - | $2\frac{1}{2}$ | - | $7\frac{1}{2}$ | - | 10 | 7 | $\frac{1}{2}$ |
| **1961 - 1971** | - | 2 | 2 | 6 | - | 10 | 7 | $\frac{1}{2}$ |
| **1972 - 1999** | $1\,\frac{19}{21}$ | - | $1\,\frac{19}{21}$ | $5\frac{15}{21}$ | $\frac{10}{21}$ | 10 | 7 | $\frac{1}{2}$ |
| **2000 - current** | $2\,\frac{6}{17}$ | - | - | $7\,\frac{1}{17}$ | $\frac{10}{17}$ | 10 | 7 | $\frac{1}{2}$ |
| | Fixed weights | | | | | | | |
| **1956 - 2010** | 2 | 2 | 2 | 6 | $\frac{1}{2}$ | 10 | 7 | $\frac{1}{2}$ |

**Weights within Tournaments**

The structure of tournaments has changed a lot over time. In general, tournaments are built up by stages, where in each stage a number of teams is eliminated. Stages are organized either as group stages or as knock-out stages. In both cases, teams play against each other either in one or two matches. Also the number of participating teams differs over the seasons and over the tournaments. For a fair comparison of player performances the following assumptions are formulated.

***Assumption 5: "Number of tournament stages"*** *The importance of a tournament is independent of the number of stages in that tournament.*

Assumption 5 yields that, for example, becoming European champion in a tournament with three stages is equally important as becoming European champion in a tournament with four stages. Correcting for the type of the stage is done based on the following assumption.

***Assumption 6: "Type of tournament stage"*** *The importance of a stage in a tournament does not depend on the type (knock-out or group) of that stage.*

Usually half of the teams is eliminated at the end of a stage. However, there are tournaments where this is not the case. For example, in the quarter finals of the FWC of 1982 the number of participants was reduced from twelve to four, while in the quarter finals of the FWC of 1986 the number of participants was reduced from eight to four. Based on this, we consider a result in the quarter finals of 1982 to be more important than a corresponding result in the quarter finals of 1986.

***Assumption 7: "Elimination in tournament stages"*** *The importance of a stage depends on the number of teams that is eliminated in that stage.*

Based on the last three assumptions, we define a weighting function that corrects for the phenomena mentioned in these assumptions. To that end, we introduce the concept of Stage Base Weight (SBW). Let $t \in T$, $y \in Y$, $i = 1, \ldots, I_{ty}$. Then the SBW is denoted and defined as,

$$b_{tyi} \quad = \quad \text{the relative importance of stage } i \text{ in tournament } t \text{ in season y.} \quad (3.6)$$

We assume an increasing trend in the values of the SBW when the tournament evolves to the next stage; i.e., a next stage is always at least as important as the previous stage. So the values of $b_{tyi}$ will satisfy:

$$b_{ty1} \leq b_{ty2} \leq \ldots \leq b_{tyI_{ty}}. \tag{3.7}$$

In the ranking model we use the following formula for calculating the values of $b_{tyi}$. It can be easily checked that $b_{tyi}$ then satisfies (3.7) . We define

$$b_{tyi} \quad = \quad \frac{1}{tyI_{ty} - i + 1}. \tag{3.8}$$

This definition means that the SBW of any final stage satisfies $b_{tyI_{ty}} = 1$. For semi-finals we have that $b_{tyI_{ty}-1} = \frac{1}{2}$, and so on for the other stages. With the SBW as starting points, we are now able to translate Assumptions 5, 6, and 7 into mathematics.

- Assumption 5: For each season and each tournament in that season the total value of all SWBs in that tournament is equal to one.
- Assumption 6: For each season and each tournament in that season, the SWB value of a certain stage is divided by the number of matches in that stage.
- Assumption 7: For each season and each tournament in that season, the SWB value of a stage is multiplied by the number of teams in that stage divided by the number of teams in the next stage.

For each $t \in T$, $y \in Y$, and $t = 1, \ldots, I_{ty}$, we define:

| | |
|---|---|
| $\rho_{tyi}$ | = the weight of stage $i$ in tournament $t$ in season $y$; |
| $a_{tyi}$ | = the number of teams in stage $i$ of tournament $t$ in season $y$; |
| $a_{ty(I_{ty}+1)}$ | = 1; |
| $J_{tyi}$ | = the total number of matches that any team has to play in stage $i$ of tournament $t$ in season $y$; |
| $w_{tyi}$ | = the weight of the matches in stage $i$ of tournament $t$ in season $y$. |

We then obtain the following formulas:

$$\rho_{tyi} = \frac{b_t yi(\frac{a_{tyi}}{a_{ty(i+1)}})}{\sum_{i=1}^{I_{ty}}(b_t yi(\frac{a_{tyi}}{a_{ty(i+1)}}))}, \tag{3.9}$$

and

$$w_{tyi} = \frac{\rho_{tyi}}{J_{tyi}}. \tag{3.10}$$

A complete Qualifiers tournament for an UEC or a FWC is considered as one stage, as we do with a SC. Hence, for a one-stage tournament $t$ in season $y$, it follows that $I_{ty} = 1$, so that $\rho_{tyi} = 1$ for each $i \in \{1, \ldots, I_{ty}\}$; the value of the match weight $w_{tyi}$ is then calculated from (3.10).

We now show that with the above definitions and formulas Assumption 4 actually holds. For all $y \in Y, t \in T$, we have that

$$\begin{aligned}
\sum_{i=1}^{I_{ty}} \rho_{tyi} &= \sum_{i=1}^{I_{ty}} \frac{b_t yi(\frac{a_{tyi}}{a_{ty(i+1)}})}{\sum_{i=1}^{I_{ty}}(b_t yi(\frac{a_{tyi}}{a_{ty(i+1)}}))} \\
&= \frac{\sum_{i=1}^{I_{ty}} b_t yi(\frac{a_{tyi}}{a_{ty(i+1)}})}{\sum_{i=1}^{I_{ty}}(b_i(\frac{a_{tyi}}{a_{ty(i+1)}}))} \\
&= 1.
\end{aligned}$$

The following example shows how the above formulas are used in our final ranking model.

**Example 2.** *The UEC of 2008 was played with sixteen teams in four stages, i.e.,* $I_{UEC-2008} = 4$. *In each stage, half of the teams was eliminated:* $a_{UEC-2008-1} = 16$, $a_{UEC-2008-2} = 8$, $a_{UEC-2008-3} = 4$, $a_{UEC-2008-4} = 2$, $a_{UEC-2008-5} = 1$. *The first stage* ($i = 1$) *was organized as a group stage, where each team had to play three matches:* $J_{EC-2008-1} = 3$. *The other stages were knock-out stages with one match:* $J_{UEC-2008-2} = J_{EC-2008-3} = J_{UEC-2008-4} = 1$. *The values of the stage weight* $\rho_{tyi}$ *and of the match weight* $w_{tyi}$ *are calculated using equations (3.9) and (3.10). For the first stage the calculations are as follows:* $\rho_{UEC-2008-1} = \frac{(\frac{1}{4})(\frac{16}{8})}{4\frac{1}{6}} = 0.12$, *and* $w_{UEC-2008-1} = \frac{0.12}{3} = 0.04$. *The results of the calculations for the other stages are listed Table 3.7.*

**Table 3.7.** Weights UEC 2008

| **Round** | $J_{tyi}$ | $b_i$ | $b_i \times \frac{a_{tyi}}{a_{ty(i+1)}}$ | $\rho_{tyi}$ | $w_{tyi}$ |
|---|---|---|---|---|---|
| First stage | 3 | $\frac{1}{4}$ | $\frac{1}{4} \times \frac{16}{8} = 0.5$ | 0.12 | 0.04 (three matches) |
| Quarter finals | 1 | $\frac{1}{3}$ | $\frac{1}{3} \times \frac{8}{4} = \frac{2}{3}$ | 0.16 | 0.16 |
| Semi finals | 1 | $\frac{1}{2}$ | $\frac{1}{2} \times \frac{4}{2} = 1$ | 0.24 | 0.24 |
| Final | 1 | 1 | $1 \times \frac{2}{1} = 2$ | 0.48 | 0.48 |
| **Sum** | 6 | $2\frac{1}{12}$ | $4\frac{1}{6}$ | 1 | |

### Field positions

The last correction we have to make concerns the field positions. According to Criterion 2 (see Section 3.3.1) no position should be in favor of another. In the next section we analyze whether or not there are differences in career length, matches played and playing time between the four field position types.

**Career Length** Usually goal keepers have longer careers than forwards. In order to analyze the length of careers in this data set, we have plotted the number of seasons players have been active in international club and international country matches. The length of the career is determined by the difference between the first season and the last season in which points are scored. As the database also contains players who are still active, this length is censored to the right, and therefore the calculated means are somewhat underestimated. The problem is that, since we measure international careers, we can not actually tell which players are censored. Some active players might never enter an international tournament anymore. But, as the censoring takes places on all four positions and the position ratio's per decade are almost equal, we may assume that not taking into account the censoring does not influence our conclusions.

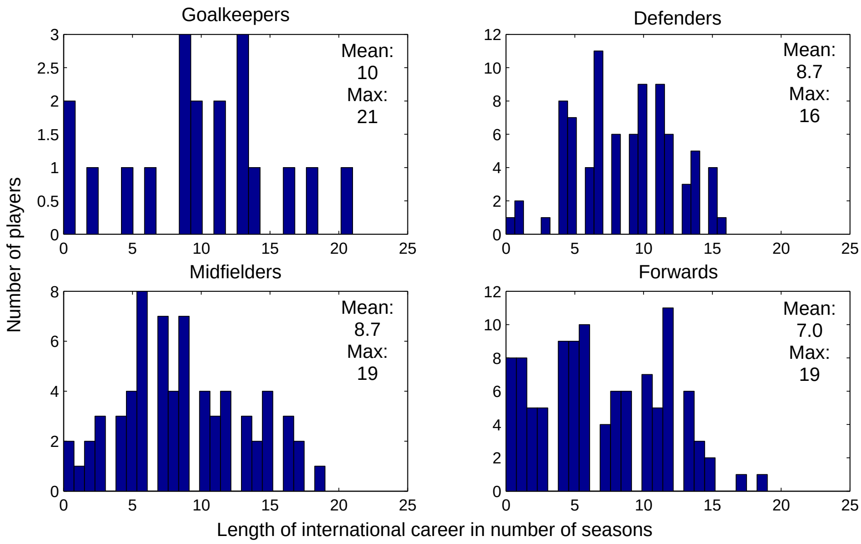



**Figure 3.1.** International career lengths

Figure 3.1 shows that goalkeepers have, on average, the longest international careers, with as absolute maximum Edwin van der Sar, who played twenty-one seasons in international matches. However, the other positions also have players with long international careers, while most other goalkeepers have less extreme careers. We used a t-test to compare the means and only found a significant difference between goalkeepers and forwards. However, when we ignore the players with five or less matches the significant difference disappeared.

The histograms from Figure 3.1 all have more or less the same shape, and show that for all positions, most players have a career length between five and ten seasons. For this ranking we assume that the average lengths of the international careers do not differ between field positions. Although we expected some differences in career lengths among field positions, with the exclusion of forward players who only played a few matches, no difference was found, so we will make no corrections here.

This counter intuitive phenomenon may be caused by the fact that, although goalkeepers retire at older ages, they also start their international career later. In many cases, goalkeepers enter the top teams around the age of twenty-four whereas talented forwards and midfielders are often lined up at much younger ages. For example, Van der Sar made his debut for Ajax when he was twenty-two, whereas Seedorf did this at the age of sixteen, (youngest debuted player ever). Since most coaches want experienced goalkeepers and hold on to their older goalkeepers. They retire at a older ages, so that the young goalkeepers often have to wait for their chances.

**Matches and minutes played** Besides the length of international careers, we also look at the number of matches and the number of minutes played. We may expect that, on average, there occur differences between positions. Coaches may change the tactics during a match, and may, for example, substitute a forward for an extra defender to preserve a lead. In this case the forward was not performing bad but is 'punished' by the tactics of the coach. In order to verify whether or not coaches use more different forward players than defenders or that forwards are more often substituted because they are tired, we summarize some statistics regarding the number of matches and the amount of playing time for all field position; see Table 3.8.

**Table 3.8.** Match statistics per field position

| | **Total** | **GK** | **DF** | **MF** | **FW** |
|---|---|---|---|---|---|
| Number in database | 274 | 19 | 82 | 67 | 106 |
| Average number of matches per player | 45 | 54 | 49 | 52 | 37 |
| Max number of matches per player | 207 | 207 | 191 | 184 | 147 |
| Average playing proportion per match | 0.90 | 0.97 | 0.93 | 0.91 | 0.85 |
| Average minutes played | 81 | 88 | 83 | 81 | 77 |

The first row of Table 3.8 shows the number of players in our data set for each of the four field positions. The second row gives the average number of matches with at least one minute played; the third row shows the maximum number of matches played by one player. In Figure 3.2 we have plotted the number of matches for each player in a histogram. The figures in 3.8 and the histograms of Figure 3.2 show little difference between goalkeepers, defenders and midfielders. Forwards play on average less matches, but only significantly less then defenders. However, as in the career length, the difference is mainly caused by the large number of forwards with few matches in our data set. Compared to the other field positions, a lot of forwards are selected for the national team, but many of them played only one or two matches. Furthermore, seven of the thirteen forwards with less than five matches are players from the starting period 1956-1960. Excluding them, increases the average to 41 matches. We conclude that no correction is needed here: On average, each player plays the same number of matches during his international career.

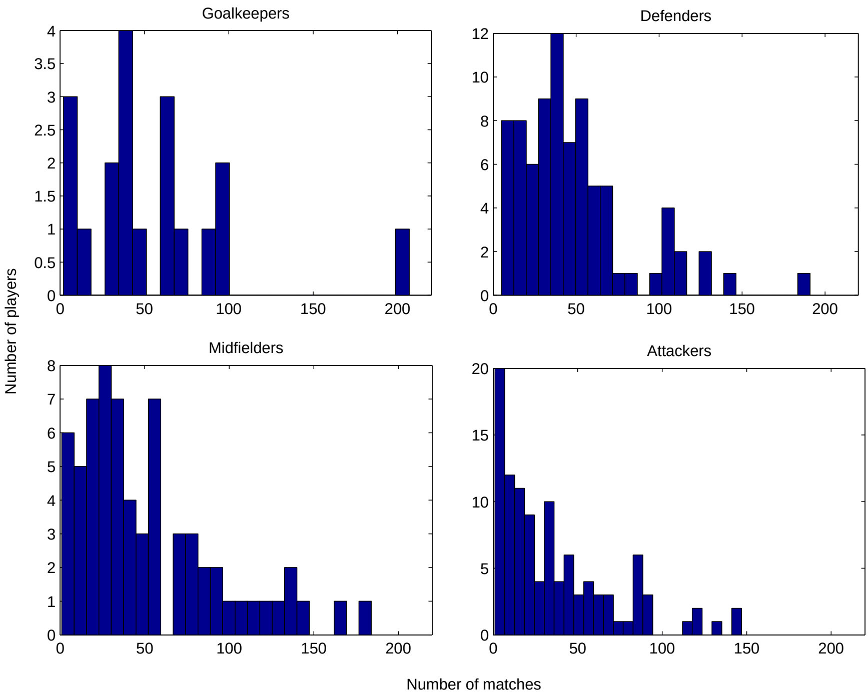



**Figure 3.2.** Number of matches per field position

Row four and five of Table 3.8 concern playing time in minute. By "Average playing proportion per match" we mean the average of the values of $\frac{m_{tyijp}}{90}$ per position. When multiplying this average by 90, we obtain the average number of minutes played per player per match per position. The figures show that goalkeepers play almost all 90 minutes, whereas forwards play on average only 77 minutes.

Since our dataset contains both matches of the national team, where all players have the Dutch nationality, and club matches, where players do not need to have the Dutch nationality, it may be interesting to look at the data from matches of the national team only. In national team matches, Dutch players are substituted by Dutch players, so that the total playing time is always 90 minutes. For club matches, a Dutch player may be replaced or may replace a non-Dutch player which is not included in the data set. Table 3.9 is similar to Table 3.8, with the exception that in the calculations of Table 3.9 only the international matches of the Dutch national team are used. As we are interested in the ratio of playing time between positions of the starting line-up we add an extra condition, namely, for the calculation of the average playtime we only use matches in which the player played at least 25% of the total match time. This means that players with only 23 minutes or less are excluded. The reason for this choice is that often offensive players are substitued by another offensive player in the last ten minutes of the match. Including these substitutes will decrease the average playing times of the offensive positions. In Section 3.5.1 also show the results if we omit this correction. In Figure 3.3 we show how the playing

**Table 3.9.** Match statistics per position of the national team

| Field position | **GK** | **DF** | **MF** | **FW** |
|---|---|---|---|---|
| Av. number of matches per player | 14 | 13 | 14 | 9 |
| Av. playing proportion per match | 0.97 | 0.93 | 0.91 | 0.85 |
| Av minutes played | 89 | 85 | 82 | 78 |
| Av. playing time relative to GK | 1 | 1.04 | 1.09 | 1.14 |

time between field positions differs. On the horizontal axis we have denoted the number of playing minutes, and on the vertical axis, per position, the percentage of matches a player plays that amount of minutes. The figure shows that during 97% of the matches goalkeepers play the full 90 minutes of the match, while only in 61% of the matches a forward reaches the final whistle. Futhermore, it shows that especially at halftime forwards are substituted. The statistics given in Table 3.9 are based on players with at least 25% playing minutes. This corresponds to the vertical line at 23 minutes. A more detailed analysis can be obtained when estimating a so-called survival curve; see, e.g., Elandt and Johnson (1980). We leave this analysis for further research.

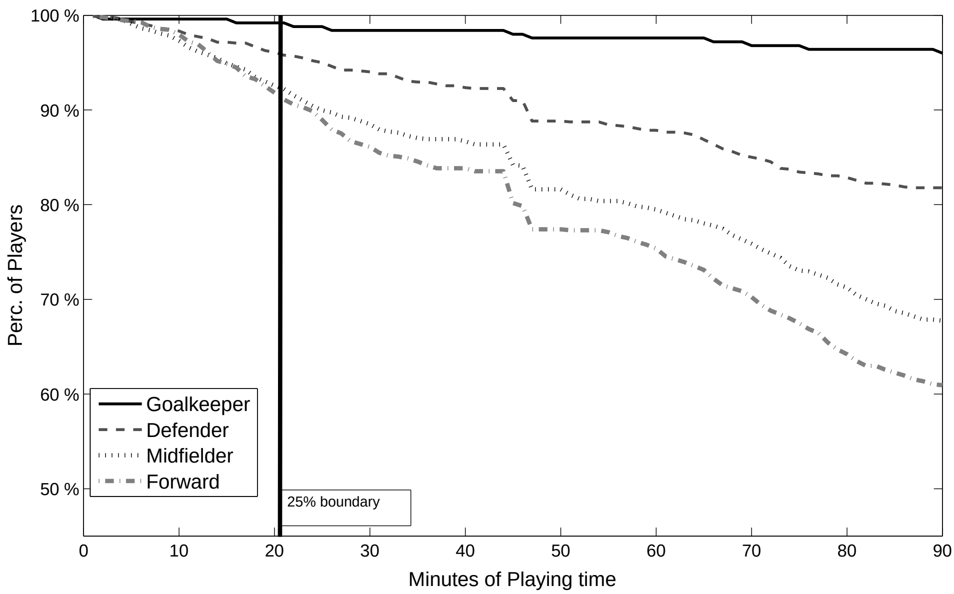



**Figure 3.3.** Playing time per field position.

There may be several reasons for the differences in Figure 3.3. Goalkeepers are in general less active than other field players and therefore these other field players are more likely to get substituted because of being exhausted or injured. In case of tactical changes of the coach, forwards are more often substituted to preserve a lead or to force a victory. Based on these arguments and the figures from Table 3.9 we make the following assumption.

***Assumption 8: "Number of minutes played"*** *The contribution to the Performance Score of the average playing time per match is the same among all field positions*

So we will only correct for the difference in playing time between field positions and not for the number of matches played or the lengths of the careers. For the correction we will use the average playing proportion compared to the average playing proportion of goalkeepers per match, given in Table 3.9. For each player $p \in P$, and each position $f \in F$ we define:

$$f_p = \text{the field position of player } p \tag{3.11}$$
$$m_f = \text{the average playing proportion per match per position } f. \tag{3.12}$$

The field position correction parameter $c_p$, is defined as:

$$c_p = \frac{m_{GK}}{m_{f_p}}. \tag{3.13}$$

The four values of $c_p$ can be found in the last row of Table 3.9. In Section 3.5.1 a ranking without a correction for the positions is compared with the ranking with corrections.

### The final ranking model

Now that we defined all assumptions we can express the equation for calculating the Performance Score. For each $i \in \mathcal{I}_{ty}, j \in \mathcal{J}_{tyi}$, $p \in P$, the *Match Score* of player $p$ in match $j$ of stage $i$ during tournament $t$ in season $y$ is denoted and defined as:

$$s_{tyijp} = 10 \times \frac{m_{tyijp}}{90} r_{tyijp} \tau_{ty} w_{tyi} c_p, \tag{3.14}$$

and the *Performance Score* of player $p$ for his complete career is denoted and defined as:

$$S(p) = \sum_{t,y,i,j} s_{tyijp}. \tag{3.15}$$

$$\tag{3.16}$$

The final ranking uses the values of $S(p)$ in ascending order; i.e., the player with the highest Performance Score is on top of the list. The complete ranking model is given in Table 3.10. The calculations are illustrated by means of Example 3. We calculate two Match Scores of Orlando Trustfull in matches of the national Dutch team.

**Example 3.**

- ***06-09-1995, Netherlands - Belarus: 1 - 0*** *This was a Qualifier for the EC of 1996. So* $t = Q$, $y = 1995/1996$, $i = 1$ *(recall that Qualifiers for one tournament are considered as a tournament with only one stage), and* $j = 8$ *(since this was the eighth match in this Qualifier tournament). Moreover,* $\tau_{Q-1995} = V_{Q-1995} = 1$. *The Netherlands needed to play eleven matches in this Qualifier tournament, so that* $J_{Q-1995-1} = 11$. *From (3.10) it follows that* $w_{Q-1995-1} = \frac{1}{11}$. *Trustfull entered the match in minute 71. Since he played during 19 minutes, it follows that* $m_{Q-1995-1-8-Trustfull} = 19$. *Since the Netherlands won this match, we have that* $r_{Q-1995-1-8-Trustfull} = 4$. *Trustful was a midfielder, so his position parameter satisfies* $c_{f_{Trustfull}} = c_{MF} = 1.18$. *The total score for this match is calculated now as follows:*

$$\begin{aligned} s_{Q-1995-1-8-Trustfull} &= 10 \times \frac{m_{Q-1995-1-8-Trustfull}}{90} \\ &\quad \times r_{Q-1995-1-8-Trustfull} \times \\ &\quad \tau_{Q-1995} \times w_{Q-1995-1} \times c_{MF} \\ &= 10 \times \frac{19}{90} \times 4 \times 1 \times \frac{1}{11} \times 1.18 \\ &= 0.91. \end{aligned}$$

- ***11-10-1995, Malta - Netherlands: 0 - 4*** *This also was a Qualifier for the EC of 1996. Therefore,* $t = Q$, $y = 1995$, $i = 1$ *and* $j = 9$ *(it was the ninth match in this Qualifier tournament). Moreover,* $\tau_{Q-1995} = 1$ *and* $w_{Q-1995-1} = \frac{1}{11}$. *Trustfull entered the match in minute 80, so he played during 10 minutes:* $m_{Q-1995-1-8-Trustfull} = 10$. *The Netherlands won this match, so that* $r_{Q-1995-1-8-Trustfull} = 4$. *The total score for this match is:*

$$\begin{aligned} s_{Q-1995-1-9-Trustfull} &= 10 \times \frac{m_{Q-1995-1-9-Trustfull}}{90} \\ &\quad \times r_{Q-1995-1-9-Trustfull} \times \\ &\quad \tau_{Q-1995} \times w_{Q-1995-1} \times c_{MF} \\ &= 10 \times \frac{10}{90} \times 4 \times 1 \times \frac{1}{11} \times 1.18 \\ &= 0.477. \end{aligned}$$

- ***Total score*** *The total score of Orlando Trustfull for the above two interland matches is* $0.9059 + 0.4768 = 1.3826$.

**Table 3.10.** The All-time Dutch soccer ranking model

| | | |
|---|---|---|
| **INPUT** | | |
| **Index sets** | | |
| $Y$ | = | $\{1956, \ldots, 2008\}$, the set of seasons; |
| $P$ | = | the set of players; |
| $T$ | = | $\{WC, EC, Q, EC1, EC2, IFC, EC3, SC\}$, the set of tournaments; |
| $\mathcal{I}_{ty}$ | = | $\{1, \ldots, I_{ty}\}$ for $t \in T, y \in Y$, the set of stages tournament $t$ in season $y$, where $I_{ty}$ denotes the number of stages in tournament $t$ in season $y$; |
| $\mathcal{J}_{tyi}$ | = | $\{1, \ldots, J_{tyi}\}$ for $t \in T, y \in Y, i \in \mathcal{I}_{ty}$, the set of matches in stage $i$ in tournament $t$ in season $y$, where $J_{tyi}$ denotes the total number of matches in stage $i$ of tournament $t$ in season $y$; |
| $F$ | = | $\{GK, DF, MF, FW\}$, set of field positions. |
| **Initial parameters** | | |
| $V_{ty}$ | = | the relative importance of tournament $t$ if organized in year $y$, 0 otherwise; |
| $c$ | = | the total number of points that can be collected from European club tournaments in one season; |
| $b_i$ | = | the base weight of stage $i$; |
| $a_{tyi}$ | = | the number of teams in stage $i$ in tournament $t$ in year $y$. |
| **Model parameters** | | |
| $\tau_{ty}$ | = | the weight factor of tournament $t$ in season $y$; |
| $\rho_{tyi}$ | = | the weight of stage $i$ of tournament $t$ in season $y$; |
| $w_{tyi}$ | = | the weight of the matches in stage $i$ in tournament $t$ in season $y$; |
| $f_p$ | = | the field position of player $p$; |
| $c_p$ | = | the field position correction for player $p$. |
| **Data** | | |
| $m_{tyijp}$ | = | the number of minutes player $p$ has played in match $j$ in stage $i$ in tournament $t$ in season $y$; |
| $r_{tyij}$ | = | the number of points for the result of the match $j$ in stage $i$ in tournament $t$ in season $y$. |
| **MODEL** | | |
| For $t \in T, y \in Y, i \in \mathcal{I}_{ty}, j \in \mathcal{J}_{tyi}, p \in P$: | | |
| $s_{tyijp}$ | = | $10 \times \frac{m_{tyijp}}{90} r_{tyijp} \tau_{ty} w_{tyi} l_p;$ |
| $S(p)$ | = | $\sum_{t,y,i,j} s_{tyijp}.$ |
| Sort the values of $S(p)$ in decreasing order. | | |
| **OUTPUT** | | |
| The all-time Dutch soccer player ranking | | |

## 3.4 Ranking results

In this section we present the results from the calculations with the model of Table 3.10 on the data set introduced in Section 3.3.2. The data is collected from several sources. The data of the Dutch national teams is collected from the websites www.voetbalstats.nl and www.weltfussball.de. The club data from the period 2001-2011 is also gathered from the website www.weltfussball.de; all other club data is gathered from old yearbooks by hand. [8]

We have encountered several problems with the data sets. Names of the players were not always correctly spelled or written in different ways (for example, Klaas-Jan Huntelaar or Klaas Jan Huntelaar) and had to be unified. Furthermore, data was missing on the websites, (some qualifiers for the UEC were missing). To ensure that we included all matches, we counted the number of matches played for both nation and club of several players and compared this with other websites, who also listed the number of matches played, for example, www.vi.nl.

In this section we present the top 20 list; the complete ranking is given in the appendix. Furthermore, we present rankings of the best international and club player, the best player per field position and the best player per decade. In Section we will discuss in particular to what extent the assumptions are met that are made in Section 3.3.2. This Section will also deal with sensitivity analysis on the parameter values. We will analyze how the rankings change if we exclude the field position corrections, use the budget weights instead of the fixed weights, and analyze the influence of changing the tournament weights.

[8]We are very gratefully to Ronald Kres, who collected all club data from the period before 2001.

### 3.4.1 Results

The top 20 (see Table 3.11) is calculated with a fixed weight for each club tournament. In Section 3.5.2 we discuss the difference between the situation of a fixed budget weight per season (Assumption 4a) and a fixed tournament weight (Assumption 4b), and why for the latter option is chosen. The column 'Period' in Table 3.11 refers to the seasons in which the player was active on the international level (of course, most of the players have longer careers). The last three columns present the total points scored, the points scored in national team matches, and in club matches, respectively.

**Table 3.11.** Top 20 of Dutch soccer players

| Rank | Name | Pos | Period | Total | Nat. | Club |
|---|---|---|---|---|---|---|
| 1 | Ruud Krol | DF | 1970-1983 | 1549 | 747 | 802 |
| 2 | Edwin van der Sar | GK | 1993-2011 | 1542 | 410 | 1132 |
| 3 | Johan Cruijff | FW | 1967-1984 | 1502 | 387 | 1115 |
| 4 | Arie Haan | MV | 1971-1984 | 1483 | 673 | 810 |
| 5 | Clarence Seedorf | MV | 1993-2011 | 1424 | 217 | 1207 |
| 6 | Johan Neeskens | MV | 1971-1981 | 1406 | 494 | 912 |
| 7 | Frank Rijkaard | DF | 1981-1995 | 1361 | 453 | 908 |
| 8 | Wim Suurbier | DF | 1967-1978 | 1338 | 393 | 945 |
| 9 | Frank de Boer | DF | 1990-2004 | 1152 | 379 | 773 |
| 10 | Ronald Koeman | DF | 1983-1997 | 1145 | 449 | 696 |
| 11 | Marco van Basten | FW | 1983-1993 | 1140 | 452 | 689 |
| 12 | Wim Jansen | MV | 1967-1982 | 1085 | 600 | 485 |
| 13 | Piet Keizer | FW | 1963-1975 | 1022 | 40 | 982 |
| 14 | Johnny Rep | FW | 1972-1986 | 991 | 627 | 364 |
| 15 | Barry Hulshoff | DF | 1967-1983 | 927 | 23 | 904 |
| 16 | Edgar Davids | MV | 1992-2007 | 921 | 320 | 602 |
| 17 | Phillip Cocu | MV | 1993-2007 | 910 | 394 | 516 |
| 18 | Mark van Bommel | MV | 2000-2011 | 903 | 330 | 572 |
| 19 | Rob Rensenbrink | FW | 1967-1980 | 902 | 565 | 337 |
| 20 | Willy van de Kerkhof | MV | 1971-1988 | 860 | 486 | 374 |

Nat. = Number of points from the national team, Club = number of points from the European club tournaments

The number one and therefore best or most successful Dutch soccer player is Ruud Krol, closely followed by Edwin van der Sar. Johan Cruijff, who is widely seen as the best Dutch soccer player, appears in position three. Both Krol and Cruijff were active in the same period and both played for AFC Ajax, with which they won the EC1 three times in a row (1971, 1972, and 1973). Cruijff scored most of his points in club tournaments. This becomes clear in Table 3.12, where we have listed the top 20 players based on only their results in the national team. Cruijff is ranked on position 21, just outside the top 20. The main reason Cruijff collected less points in national matches is because he did not play the 1978 FWC, whereas Krol played and lost his second world cup final.

The number two of the list is Edwin van der Sar, a goalkeeper with an impressive career: even at the end of his complete career he played for a top club, namely Manchester United. Already in 1995 Van der Sar won his first EC1 with AFC Ajax, while thirteen years later, in 2008, he won his second one with Manchester United. Van der Sar ended his career at the age of forty in the EC1 final of 2010 and became the oldest player ever appeared in an EC1 final. Manchester lost this finale with 1-3 against FC Barcelona. By ending his international career already in 2008, he missed the FWC in 2010. In case he would have been at the 2010 FWC, with no doubt, he would have been the number one in the final ranking. Two players from the top 20 are currently still active, namely Clarence Seedorf and Mark van Bommel.

The top 20 of Table 3.12 contains quite a few players from the generation of 1974-1978. The complete top 6 is from this period and all six became two times world vice-champion. The number eight till twelve are the best players of the generation of 1988 who became European champion in that year. The current generation, that played the FWC final in 2010 is ranked just outside the top 20. A good performance in the UEC of 2012 will certainly bring some of the current generation in the top 20.

**Table 3.12.** Top 20: Dutch national team players

| Pos | Name | Total | FWC | UEC | Q |
|---|---|---|---|---|---|
| 1 | Ruud Krol | 747 | 502 | 165 | 80 |
| 2 | Arie Haan | 673 | 521 | 126 | 26 |
| 3 | Johnny Rep | 627 | 505 | 85 | 38 |
| 4 | Wim Jansen | 600 | 516 | 26 | 58 |
| 5 | Rob Rensenbrink | 565 | 494 | 27 | 44 |
| 6 | Johan Neeskens | 494 | 419 | 26 | 50 |
| 7 | Willy van de Kerkhof | 486 | 240 | 173 | 73 |
| 8 | Ruud Gullit | 458 | 15 | 373 | 71 |
| 9 | Jan Wouters | 458 | 39 | 362 | 57 |
| 10 | Frank Rijkaard | 453 | 56 | 345 | 51 |
| 11 | Marco van Basten | 452 | 15 | 371 | 66 |
| 12 | Ronald Koeman | 449 | 64 | 345 | 39 |
| 13 | Giovanni van Bronckhorst | 449 | 266 | 126 | 57 |
| 14 | Dennis Bergkamp | 428 | 205 | 169 | 53 |
| 15 | Hans van Breukelen | 413 | 13 | 329 | 71 |
| 16 | Edwin van der Sar | 410 | 155 | 167 | 88 |
| 17 | Adri van Tiggelen | 407 | 14 | 345 | 48 |
| 18 | René van de Kerkhof | 407 | 244 | 117 | 46 |
| 19 | Phillip Cocu | 394 | 170 | 167 | 58 |
| 20 | Wim Suurbier | 393 | 304 | 25 | 64 |

In Table 3.13 we have ranked the players based on the points scored in club matches. Best club player is Clarence Seedorf, who won the EC1 four times with three different clubs, and lost the final once in 2005. The number two, Edwin van der Sar won twice the EC1 and played the final in total five times. Johan Cruijff won the EC1 three times with Ajax and is ranked third. Eight players in the top 20 are players that won EC1 with AFC Ajax in 1995 for the first time, with exception of Rijkaard who already had won it with AC Milan. The other seven players were still at the beginning of their career and continued their successes at other clubs outside the Netherlands.

**Table 3.13.** Top 20: European club tournaments

| Pos | Naam | Totaal | EC1 | EC2 | EC3 | SC |
|---|---|---|---|---|---|---|
| 1 | Clarence Seedorf | 1207 | 1145 | 6 | 34 | 22 |
| 2 | Edwin van der Sar | 1132 | 1051 | 18 | 46 | 18 |
| 3 | Johan Cruijff | 1115 | 963 | 0 | 130 | 21 |
| 4 | Piet Keizer | 982 | 912 | 0 | 36 | 34 |
| 5 | Wim Suurbier | 945 | 853 | 0 | 58 | 34 |
| 6 | Johan Neeskens | 912 | 742 | 77 | 67 | 25 |
| 7 | Frank Rijkaard | 908 | 777 | 97 | 7 | 28 |
| 8 | Barry Hulshoff | 904 | 820 | 0 | 50 | 34 |
| 9 | Arie Haan | 810 | 525 | 212 | 26 | 47 |
| 10 | Ruud Krol | 802 | 688 | 0 | 80 | 34 |
| 11 | Frank de Boer | 773 | 611 | 18 | 128 | 16 |
| 12 | Gerrie Mühren | 768 | 670 | 0 | 63 | 36 |
| 13 | Sjaak Swart | 759 | 709 | 7 | 37 | 7 |
| 14 | Ronald Koeman | 696 | 606 | 38 | 18 | 34 |
| 15 | Marco van Basten | 689 | 576 | 83 | 12 | 17 |
| 16 | Edgar Davids | 602 | 537 | 9 | 44 | 11 |
| 17 | Danny Blind | 580 | 363 | 85 | 112 | 21 |
| 18 | Mark van Bommel | 572 | 498 | 0 | 74 | 0 |
| 19 | Ronald de Boer | 538 | 479 | 17 | 36 | 6 |
| 20 | Michael Reiziger | 528 | 476 | 0 | 34 | 18 |

In Table 3.14 we present the top 10 for each field position. Interesting fact is that each field position is represented in the top 4. The top 10 list per field position also shows why most people consider Johan Cruijff as the best Dutch player of all times; forward players score goals and are most remembered for that. The only active players in Table 3.14 are Mark van Bommel and Clarence Seedorf; Edwin van der Sar and Giovanni van Bronckhorst retired in 2011 and 2010, respectively. Dirk Kuijt is the highest ranked active forward, namely on the 11th position. Finally, in Table A.16 of the Appendix, we have split the ranking per decade, starting with the period 1956-1959. The last period contains only the year 2011. The period 2000-2010 who are currently the best and most successful players.

In Figure 3.4 we show level of success of the Dutch soccer in international team and club matches through the years. For each season we have added the scores of the players active in that season and plotted it against the season. The figure contains two graphs: the bold graph reflects the total score of both international as club points, while the dotted graph represents the scores from only the national team. The season 1987/1988 was a very successful year for the Netherlands which won the UEC for national teams and the EC1 (won by PSV). The three years in which the Netherlands became semi World Champion, 1974, 1978 and 2010, also show a high peak. Note that in the periods 1956-1968 and 1980-1986, the Netherlands were not that successful.

**Table 3.14.** Top 10 of Dutch soccer players per field position

| | **Top 10 Goalkeepers** | | **Top 10 Defenders** |
|---|---|---|---|
| 1 | Edwin van der Sar | 1 | Ruud Krol |
| 2 | Hans van Breukelen | 2 | Frank Rijkaard |
| 3 | Jan Jongbloed | 3 | Wim Suurbier |
| 4 | Piet Schrijvers | 4 | Frank de Boer |
| 5 | Eddy Treijtel | 5 | Ronald Koeman |
| 6 | Maarten Stekelenburg | 6 | Barry Hulshoff |
| 7 | Ed de Goeij | 7 | Jaap Stam |
| 8 | Jan van Beveren | 8 | Giovanni van Bronckhorst |
| 9 | Eddy Pieters Graafland | 9 | Michael Reiziger |
| 10 | Stanley Menzo | 10 | Berry van Aerle |
| | **Top 10 Midfielders** | | **Top 10 Forwards** |
| 1 | Arie Haan | 1 | Johan Cruijff |
| 2 | Clarence Seedorf | 2 | Marco van Basten |
| 3 | Johan Neeskens | 3 | Piet Keizer |
| 4 | Wim Jansen | 4 | Johnny Rep |
| 5 | Edgar Davids | 5 | Rob Rensenbrink |
| 6 | Phillip Cocu | 6 | Ruud Gullit |
| 7 | Mark van Bommel | 7 | Dennis Bergkamp |
| 8 | Willy van de Kerkhof | 8 | Sjaak Swart |
| 9 | Gerrie Mühren | 9 | Patrick Kluivert |
| 10 | Wim van Hanegem | 10 | Marc Overmars |

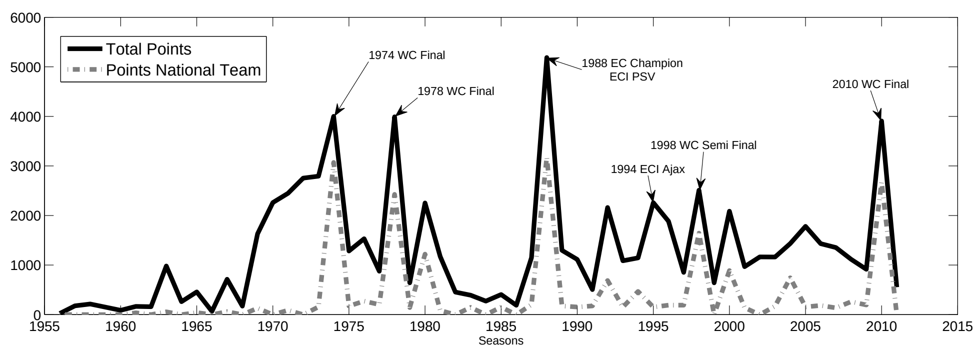



**Figure 3.4.** Performance levels over the years

## 3.5 Discussion

In Section 3.3.1 we have formulated a variety of pre-conditions to be satisfied by the final ranking. The first one, *Players from the same time period should be ranked according to the ranking in that period* can be verified by Table A.16. In this table the players are ranked per decade, while the players which were most successful in that period are in the top.

The second pre-condition, *In the top of the ranking the distribution of field positions should be more or less according to the ratio of the most common playing formation (4-3-3 or 4-4-2.)*, is analyzed as follows. The most used playing formation used by the Dutch national team is the 4-3-3 system. Therefore, we expect that in the top 100 (in total we ranked 275 players), the field positions are more or less represented and distributed by this ratio. Table 3.15 lists the expected numbers for the field positions in case of the 1-4-3-3 system. In Table 3.15 can be seen that we expect about 9 goalkeepers, 36 defenders, 27 midfielders, and 27 forwards in the top 100.

**Table 3.15.** Ratio of field position in top 100

| **Field position** | **System** | **Ratio** | **Expected in top 100** |
|---|---|---|---|
| GK | 1 | 1/11 | 9.1 |
| DF | 4 | 4/11 | 36.4 |
| MF | 3 | 3/11 | 27.3 |
| FW | 3 | 3/11 | 27.3 |

Furthermore, we do not only expect that the field positions are represented by the ratio 1-4-3-3, we also expect that they are more are less uniformly distributed over the first hundred. We verify the distribution by calculating the average ranking position, the total ranking score and the standard deviation of the ranking score. The total ranking score for each field position $f$ is calculated by adding up all ranking scores of player $i$ with field position $f$ in the top 100.

The ranking score of player $i$ is denoted and defined by

$$RS_i = 101 - rank_i$$

with $rank_i$ being the ranking position of player i in the final ranking. So the first ranked player in the final ranking has an ranking score equal to $RS_1 = 100$ and the number 100 has $RS_{100} = 1$.

The expected value for the average ranking position per field position in the top hundred of course 50.5. For the total rank score per field position we expect 50.5 times the corresponding expected number of players given in Table 3.15. We have also calculated the expected total rank score if the exact number of players per field positions in the top 100 are used. Finally the standard deviation of the ranking score per field position is given and we apply a Kruskal-Wallis test with the Null hypothesis that all positions have equal medians.

**Table 3.16.** Numbers and scores per field position in top 100

| **Statistic** | **GK** | **DF** | **MF** | **FW** |
|---|---|---|---|---|
| Expected in top 100 | 9.1 | 36.4 | 27.3 | 27.3 |
| Actual in top 100 | 9 | 33 | 28 | 30 |
| Average rank in top 100 | 52.7 | 51.9 | 47.5 | 51.2 |
| Expected total points, 4-3-3 | 459 | 1836 | 1377 | 1377 |
| Expected total points, top 100 | 454 | 1666 | 1414 | 1515 |
| Total rank points | 435 | 1621 | 1499 | 1495 |
| Standard deviation | 26.0 | 28.7 | 32.2 | 28.2 |
| | Chi-sq | p-prob | | |
| Kruskal Wallis test | 0.48 | 0.92 | | |

In Table 3.16 the results of each statistic is presented. The table shows that the field positions are represented accordingly to the 4-3-3 system. The top 100 contains slightly less defenders (33, expected 36) and some more forwards (30, expected 27). These differences can be explained by the fact that the database contains 24 more forwards than defenders, (see Table 3.8). Note that in the top 11 there are one goalkeeper, five defenders, three midfielders, and only two forwards.

The row 'Average rank in top 100' of Table 3.16 shows that the average rank position of each field position lies around 50.5. On average, midfielders are ranked highest. The row "Expected total points, 4-3-3 ' shows the expected number of rank points based on the 4-3-3 line up (so for forwards we multiply 50.5 with 27.3) and row 'Expected total points, top 100' gives the expected number of rank points based on the actual number of players in the top 100 (so for forwards we multiply 50.5 with 30). The total rank point in the last row represents the actual scores and shows us that in comparison to the expected total points 4-3-3, the defenders are a little under ranked with respect to the forwards. However, we have already seen that the number of defenders in the top 100 is lower as expected. If the actual number of players per field positions is used, the expected total points top 100 is far more comparable with the total rank points.

The standard deviation of the ranking score per field position shows that the midfielders have the highest spread, and goalkeepers the lowest. In Table 3.17 we divided the top 100 in five section and show the distribution per section, in order to clarify the spread among the top 100 more clearly. Finally the Kruskal Wallis test shows that there is no significant difference in the median of the ranking positions. The table explains the slightly higher standard deviation for midfielders as they are represented more in the positions 1-20 and 81-100. Furthermore, the goalkeepers are most found in the positions 41-60, whereas in the positions 61-80 the defenders are over represented. However, each section contains at least one goalkeeper, five defenders, three midfielders and five forwards and compared to the expected is satisfying.

From Table 3.16 and Table 3.17 we may conclude that within the top 100 the

**Table 3.17.** Distribution of field position in top 100

| Field position | GK | DF | MF | FW |
|---|---|---|---|---|
| Expected | 1.8 | 6.6 | 5.6 | 6 |
| Position 1-20 | 1 | 6 | 8 | 5 |
| Position 21-40 | 1 | 5 | 6 | 8 |
| Position 41-60 | 3 | 7 | 4 | 6 |
| Position 61-80 | 2 | 10 | 3 | 5 |
| Position 81-100 | 2 | 5 | 7 | 6 |
| Total | 9 | 33 | 28 | 30 |

field positions are pretty well represented by ratio 4-3-3 and more are less uniformly among the top 100.

### 3.5.1 Influence of field position corrections

In Section 3.3.2 we have decided to correct for the difference in playing time per match per field position. In order to analyze the influence of the correction we also run the model without the correction, i.e. $l_p = 1$ for all field positions $p$. In Table 3.18 the top 11 is given. the table shows that the same eleven players are in the top 11, but

**Table 3.18.** Top 11 of Dutch soccer players without field position correction

| Rank | Name | Pos | Period | Total | Nat | Club |
|---|---|---|---|---|---|---|
| 1 | Edwin van der Sar | GK | 1993-2011 | 1542 | 410 | 1132 |
| 2 | Ruud Krol | DF | 1970-1983 | 1475 | 711 | 764 |
| 3 | Arie Haan | MF | 1971-1984 | 1348 | 612 | 736 |
| 4 | Johan Cruijff | FW | 1967-1984 | 1317 | 339 | 978 |
| 5 | Frank Rijkaard | DF | 1981-1995 | 1296 | 431 | 865 |
| 6 | Clarence Seedorf | MF | 1993-2011 | 1294 | 197 | 1097 |
| 7 | Johan Neeskens | MF | 1971-1981 | 1278 | 449 | 829 |
| 8 | Wim Suurbier | DF | 1967-1978 | 1274 | 374 | 900 |
| 9 | Frank de Boer | DF | 1990-2004 | 1097 | 361 | 736 |
| 10 | Ronald Koeman | DF | 1983-1997 | 1091 | 427 | 663 |
| 11 | Marco van Basten | FW | 1983-1993 | 1000 | 396 | 604 |

appear in a different order. In this ranking Van der Sar is ranked first. Goalkeepers receive no field position correction in the model with correction and therefore they get the same amount of points as in the model with correction. As defenders did get a field correction, van der Sar now surpasses Krol.

Although the field position correction has not a lot of influence on the final ranking, the following arguments are used to keep the correction. First of all, still based on the argument, from Section 3.3.2, forward players play on average less minutes per match. Secondly, if we look at the number of players per field position in the top 100, see Table 3.19, an increase of two goalkeepers is seen. This means that 58% of the goalkeepers of the database are also in the Top 100. Comparing this percentage with the other field positions, it is better distributed with the field correction, although in both cases the forward position are somewhat underrepresented. However, keep in mind, that our database contains a lot of forward players and a part of them are players with only a few matches. Among them are players from the early period 1956-1960 or players used as substitute in the Dutch national team for a few matches. For example, former soccer head coach, Marco van Basten used 49 players, of which 32 were debutants, in 34 match and most of these debutants were forward players. These debutants are in the bottom of the ranking and decrease the percentage of forward players in the top 100.

**Table 3.19.** Top 100 with and without field position correction (fpc)

| **Statistic** | **GK** | **DF** | **MF** | **FW** |
|---|---|---|---|---|
| Top 100, with fpc | 9 | 33 | 28 | 30 |
| Top 100, without fpc | 11 | 34 | 26 | 29 |
| All Players | 19 | 82 | 67 | 106 |
| Percentage of total | | | | |
| Top 100, with fpc | 47% | 40% | 42% | 28% |
| Top 100, without fpc | 58% | 41% | 39% | 27% |

### 3.5.2 Fixed season budget for the European club tournament

In Section 3.3.2 we have presented two methods to assign points to European club tournaments. So far we have used the 'fixed weight' variant, where each club tournament weight has a fixed weight; see Table 3.6. One could argue that, since the number of tournaments has changed, in seasons with less tournaments it is more difficult to win a price and therefore tournaments in these seasons should be rewarded with a higher weight. This argument is used in the so called 'budget' variant, where each season a fixed amount of points (the budget), which is divided among all European tournaments according to the relative weight factors form Table 3.6. The results are presented in Table 3.20.

**Table 3.20.** Top 11: 'Seasonal budget' option for European club tournaments

| **Rank** | **Name** | **Pos** | **Period** | **Total** | **Nat** | **Club** |
|---|---|---|---|---|---|---|
| 1 | Edwin van der Sar | GK | 1993-2011 | 1635 | 410 | 1225 |
| 2 | Clarence Seedorf | MF | 1993-2011 | 1539 | 217 | 1322 |
| 3 | Ruud Krol | DF | 1970-1983 | 1516 | 747 | 769 |
| 4 | Johan Cruijff | FW | 1967-1984 | 1470 | 387 | 1083 |
| 5 | Arie Haan | MF | 1971-1984 | 1447 | 673 | 774 |
| 6 | Johan Neeskens | MF | 1971-1981 | 1373 | 494 | 879 |
| 7 | Frank Rijkaard | DF | 1981-1995 | 1318 | 453 | 865 |
| 8 | Wim Suurbier | DF | 1967-1978 | 1313 | 393 | 921 |
| 9 | Frank de Boer | DF | 1990-2004 | 1174 | 379 | 795 |
| 10 | Ronald Koeman | DF | 1983-1997 | 1112 | 449 | 663 |
| 11 | Marco van Basten | FW | 1983-1993 | 1107 | 452 | 656 |

The largest change is observed in the number of points rewarded to the EC1: in the fixed weight model the relative weight is 6, whereas in the budget version it varies from $5\frac{15}{21}$ in the period 1971-1999 to $7\frac{1}{17}$ in the period 2000-current. The differences between both options is illustrated in Figure 3.5. The vertical axis of this figure contains the difference in points between the 'budget' option and the 'fixed weight' option. It shows that in the period 1971-1999 players lose points, with the highest loss, namely 43, for Frank Rijkaard. The players active after 1999 earn more points in the budget version. Clarence Seedorf benefits the most; he increases his total score with 115 points.

When looking at the final ranking we observe that within the top 11, only the top 5 change when we use the budget version instead of the fixed. In the top 50 Mark van Bommel is the player who benefits most. He climbs from position 18 to 13. All other changes are restricted t not more than a change of three positions

The question remains whether or not it is fair to reward the EC1 with one more point because of the fact that the EC2 has disappeared. After this change in 1999, the number of participants in the EC3 has increased, while after the group stage (stage 1) in the EC1, teams, who ended third in their group, enter the EC3 in the second stage, giving them a new opportunity to gather points and win a price. The difference

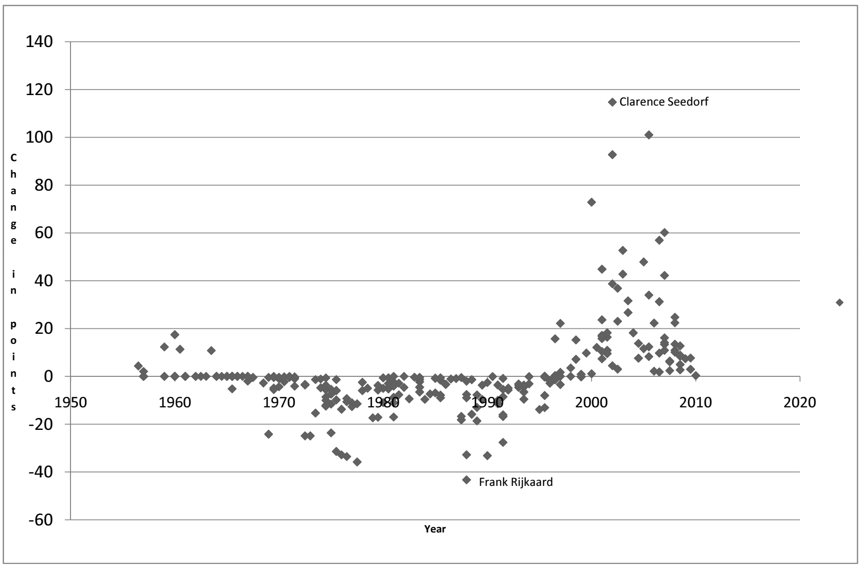



**Figure 3.5.** Differences between 'Season Budget' or 'Fixed' weight

between the $5\frac{15}{21}$ and $7\frac{1}{17}$ can be considered to be too large, two players still active after 1999, namely Van der Sar and Seedorf, occupy the first and second place in case we use the model with the budget for European club tournaments. Van de Sar and

Seedorf gain respectively 70% and 81% of their points from the EC1 and in case of the budget variant they beat all the players who were also successful in the National team because the EC1 is worth much more points. We decided to use fixed points for each European tournament for the final ranking based on this analyse. In next Section we will change the relative tournaments weights and see who in both fixed and budget version the ranking will change.

### 3.5.3 Tournament weights

Table 3.6 contains the relative tournaments weights for both fixed weights and budget weight for European tournaments. The FWC is set to be the most important tournament with a weight of 10. The UEC follows with a weight of 7. The number one of the final ranking, Ruud Krol is after Arie Haan, Johnny Rep, and Win Jansen the number four in the FWC ranking; see Table 3.12. Frank Rijkaard is the highest ranked player who won the UEC in 1988 and the first player on the ranking who has much more UEC points than Krol. The question is how much do we have to increase the relative weight of the UEC to get Rijkaard as the number one.

An increase in weight of the UEC will increase the score of all UEC players, while players with the most UEC points will benefit the most. The difference in UEC points between Krol and Rijkaard is 180 points. This is equal to (180/7=) 25.7 points per relative UEC point. The difference in total points between the two players is 188. So, if we increase the relative weight by (188/25.7 =) 7.3, Rijkaard will be ranked first. This means that we can vary the relative UEC weight of the UEC from 0 through 14.3 without harming the first position of Krol.

A similar analysis is made for the EC1, the EC2 and the EC3; see Table 3.21. The differences in points of the Q matches and SC's between all players are too small. We also looked at how much the FWC weight has to decrease in order to get Ruud Krol from the first place.

In Table 3.21 it is seen that the FWC and the EC1 value have to change only by a small value to make Van der Sar the number one: the difference between the numbers one and two is only 7 points. The table also shows us that in case of the change in FWC value and EC1 value, in the top 50 only twelve (50*0.24) players

**Table 3.21.** Winner stability with repect to changes of the $V_{ty}$ value

| Tournament | New number one | $V_{ty}$ change | new $V_{ty}$ value | Max. t50 | Avg. t50 |
|---|---|---|---|---|---|
| UEF | Frank Rijkaard | 7.3 (104%) | 14.3 | 19.0 | 5 |
| EC1 | Edwin van der Sar | 0.11 (2%) | 6.11 | 1 | 0.2 |
| EC2 | Arie Haan | 0.62 (31%) | 2.62 | 5 | 0.8 |
| EC3 | Johan Cruijff | 1.88 (94%) | 3.89 | 9 | 2.5 |
| FWC | Edwin van der Sar | -0.2 (2%) | 9.8 | 1 | 0.2 |

Max. t50: Maximum change in positions in top 50, Avg t50: Average change in positions in top 50 change one position. In the other cases there are some bigger changes in the top 50. The biggest difference can be observed when the UEF value is doubled. All UEF winners of 1988 will gain six through a maximum of nineteen positions (Jan wouters rises from the 41th place to the 22th place) whereas on average the change in positions in the top 50 is five positions. In case we change the values of the EC2 and the EC3 the changes in the top 50 are smaller. The maximum change in positions is respectively five and nine places and the average change 0.8 and 2.5 places. Still in all five rankings the top 11 remains the same.

### 3.5.4 Dutch national team: Experience level and performance

In this section we present another use of player's scores. Recall that players with a high PP-score can be considered as highly experienced, because they have many playing minutes in important matches. But is the converse also true? Can we expect that teams with many experienced players will be successful in a next tournament? Or, to what extend is the total PP-score of a selection indicative for success in a next tournament?

The above formulated question will not be studied in full detail in this section. We only compare the experience level of the Dutch national team with its performance during European and World Championships. Our definition of experience levelis directly derived from PP-scores, namely, the experience level of a team/selection at a certain time instance is the sum of the PP-scores of the players in that team/selection.

In Figure 3.6 we have depicted the experience level of the Dutch national team each time just before the start of a European or World Championship in the period 1974 through 2012. The horizontal axis in Figure3.6 refers to the years of the major tournaments, and the vertical axis contains the scale of the experience level. The names of the head coaches are written in a box just underneath the horizontal axis; the location of the coach names in the box corresponds to the years they were in function.

The graph of Figure 3.6 shows an absolute maximum in 1978, the year of the World Championship in Argentina. The Dutch selection was already pretty experienced in 1974, the year of the World Cup in West-Germany. In both years the Dutch team made it to the final, but lost in both cases from the organizing country. Johan Cruijff, considered to be the best soccer player of that time, played in West-Germany, but not in Argentina. So, the maximum in 1978 would have been even higher if Cruijff was part of the selection.

The high level of experience in the period 1974-1978 is mainly caused by players of Feyenoord and Ajax. These players won one or more European Cups I, namely with Feyenoord in 1970, and with Ajax in 1972, 1973, and 1974.

The years after the Cruijff-Krol period can be characterized as the lean years of Dutch soccer: for three consecutive tournaments the Dutch national team did not qualify for the big tournaments.

In 1988, the new head coach Rinus Michels entered the European Championships

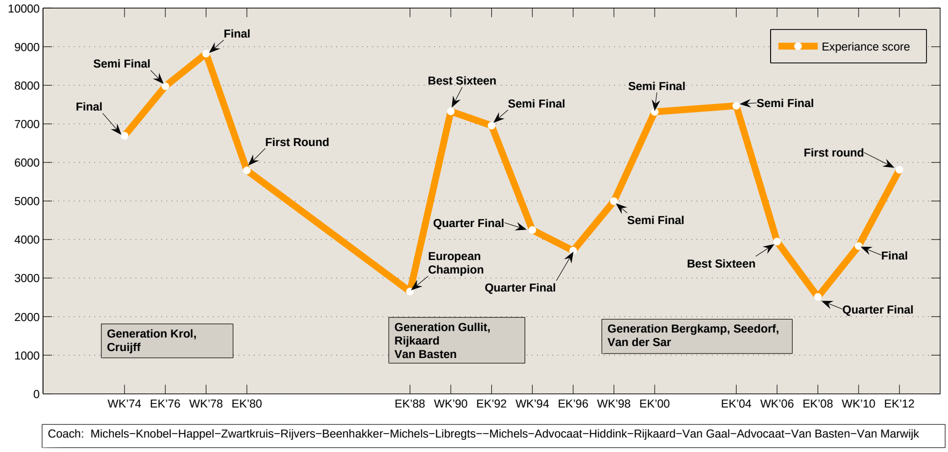



**Figure 3.6.** Experience level of Dutch national team

in West-Germany with a relatively inexperienced selection, actually a selection with a lowest total PP-score in the history of Dutch professional soccer. Players like Ruud Gullit, Marco van Basten, and Ronald Koeman were at the beginning of their careers. However, with a fabulous goal, Van Basten shoots the Netherlands to its first big victory. No wonder that after this success most players were transferred to the major European clubs, with which they also won major prices. Nevertheless, this extra experience did not lead to a second success during the World Championship in Italy in 1992: already in the group phase the team performed far below expectations, and did not even manage to proceed to the quarter finals.

In 1994, a new generation arrived at the scene. The core of this selection consisted of the Ajax players Edwin van der Sar, Clarence Seedorf, Patrick Kluivert, Edgar Davids, and Frank and Ronald de Boer. With these players, Ajax won the Champions League in 1994. In 2000, this generation seemed to be ready for big victories. However, they only made it three times to a semi-final, and are known since then as the second golden generation without gold.

In 2006, Marco van Basten started building a new team with Van der Sar and Philip Cocu as a backbone. This relatively young selection played a number of strong matches during the tournaments in 2006 and 2008, but could never reach the absolute top. Van Basten resigned and the new head coach Bert van Marwijk could profit from the work of Van Basten and the fact that in the mean time most players were active and successful in the major European teams. Already at his first tournament Van Marwijk and his players reached the World Cup final, but for the third time in its history, the Netherlands became second.

Due to this performance the expectations for the next European Championships in 2012 were of course high. The graph of Figure 3.6 shows a pretty high experience level, but again this was not a guarantee for success. On the contrary, the Dutch team ended just above Ireland, who ended last. All three matches in the group stage were lost and Van Marwijk resigned directly after the deception.

## 3.6 Conclusion

Also in sports one wants to know: Who was or is the best player or team? In this chapter we have formulated a method for comparing Dutch soccer players, active since the season 1955/1956. We have based the performance indicators on results obtained in the international arena, namely from international club matches and matches of the national team during the big tournaments. We used match result, and assumed that team members with an equal number of playing minutes contribute equally to the result of the match.

Furthermore, we took into account the importance of the tournament, and the stage of the tournament. We also distinguished between the four types of field positions. We have calculated the average playing times for the players in these field positions. In our model, these differences were taken into account by applying corrections on the various playing times.

The final ranking shows as top three: 1: Ruud Krol, 2: Edwin van der Sar, and 3: Johan Cruijff. The difference between the number one and two is rather small and sensitivity analysis shows that small changes in the tournament weights and position corrections may switch the positions one and two. Further analysis shows that the top eleven of the ranking is rather insensitive for the various sensitivity scenarios.

The logical next research step is to include non-Dutch players. Collecting all the necessary data will be a major bottleneck in this research. Currently more and more data is stored on the internet and in databases. So in case of the last two decades, the collection is less hard and for this period a worldwide ranking could be made. But if it is possible to collect all data, would it not be great to compare the big players, like Pele, Maradona and Beckenbauer, and see who was the most successful player of all times?

Chapter 4

# Selection Procedures for Olympic Speed Skating

Selecting the best athletes for major tournaments is usually a controversial affair, especially when there are more candidates than starting positions. The Netherlands faces this problem for speed skating at Olympic Winter Games. The problem is even more difficult, because some skaters need to start at more than one distance, while other skaters may have a much higher win probability on one of these distances. The reason is that the Netherlands has eighteen starting positions, but can delegate only ten of its best skaters. This holds for both male and female skaters. This paper presents a binary linear optimization model with which a team of skaters can be selected with a highest probability of winning medals. The win probabilities are based on results from pre-seasonal tournaments. The chapter also describes how the Royal Dutch Speed Skating Union (KNSB) has used the results for their final decisions for the 2010 Winter Olympics in Vancouver, Canada.

## 4.1 Introduction

The Olympic Winter Games is one of the major sport events in the world. It is organized each four years. In 2010, the twenty-first Olympic Winter Games took place in Vancouver, Canada. Speed skating is one of the events that has been present since the first Winter Olympics in 1924. For the Netherlands, speed skating is the most important discipline during the Winter Games: before 2010, twenty-two of the twenty-five Dutch golden medals are won in speed skating [9]. So the success during Winter Games is completely determined by success in speed skating. However, since the the number of Dutch skaters that meet the Olympic criteria exceeds the number of available positions, a selection procedure is necessary. Of course, this procedure matters a big deal to both the athletes and the Royal Dutch Speed Skating Union (KNSB); see KNSB (2010).

The KNSB wants of course to select its best athletes and win as many (gold) medals as possible. However, such a goal is not easy to achieve, as the selection procedure itself may lead to unwanted choices. Should the procedure be based on a single selection event or on a number of tournaments, and if is chosen for the latter, should one take the best performance or an average performance as selection criterion? Hizen and Okui (2009) describe and analyze three selection procedures, namely a single race procedure, and two procedures with multiple races, one uses best results, and the other average results. Based on game theoretical strategies this paper reveals incentives of athletes, such as, whether or not to compete in a second match, and it shows the procedure which selects with the highest probability the athlete with the highest ability. In Ryvkin (2010) such athlete strategies are not taken into account, but a best selection method is chosen by looking at the efficiency of the procedure. For three tournament set-ups, namely single event, binary elimination and round robin, this paper tests the efficiency of the procedures based on the expected ability of the winner and the expected rank of the winner.

The situation for the KNSB is more complicated than the situations described in the literature above. The KNSB has to select four male skaters for the distances 500m, 1000m and 1500m, and three for the 5000m and the 10000m. For female skaters the situation is the same, except that they skate the 3000m instead of the 10000m, and have only two places at the 5000m; see Section 4.3.2. Moreover, both for male and female skaters there is a maximum of ten skaters in total. Hence, some skaters have to start on more than one distance in order to fill al eighteen places. This means that not only skaters within one discipline need to be compared, but also skaters from different disciplines. This yields a second complication, as it is hard to compare results from different distances. So we need to design a criterion with which such athletes can be compared, and that enables us to compare results from, for example, a 500m skater with a 10000m skater.

Selection tournaments are usually not very popular with athletes as they may interfere with their training programs. An athlete wants to peak at the most important tournament and his complete training schedule of sometimes multiple years is tuned

[9]Before 2010, the other three are won in figure skating

at this event. Training programs are often split up in periods of endurance training, intensive training and rest, following each other up in cycles. These cycles of a periodization training program are fit in such a way that the athlete can peak at the right time; see, e.g., Kraemer and Fleck (2007). Selection tournaments may disturb training programs in multiple ways. First, the training schedule need be adjusted in a way such that the selection tournament fits into the training cycles. Secondly, an extra peak moment needs to be scheduled in the program as the athlete has to perform well enough at the selection tournament. The occurrence of another peak moment may come at the cost of training time and may reduce the performance later on. The athlete has to choose how much training time he wants to loose in order to perform well at the selection tournament. However, losing too much time will reduce his chances on the major tournament, but too little time may result in no selection at all.

Therefore, the timing of a selection tournament is important as it should not interfere too much with the skater's preparation. From the athlete point of view, the selection should take place at least a year before the major tournament. For the KNSB, the final selection maker, it is more preferable to organize the selection tournament more close to the major tournament as they want to sent the best athletes of that moment and not the best athletes of a year ago. In such a complicated selection process all these aspects have to be taken in account.

Before 2009, the KNSB based its selection on a skate-off tournament, called the Olympic Qualification Tournament (OQT), at the end of December, two months before the Olympics. Skaters could take part in this tournament, if they had received a so called Olympic Nomination: a top-eight classification during one of the recent World Cups; see Section 4.4. During the OQT, nominated skaters, who finished within the first two on a distance, were selected automatically. In theory, ten skaters could have been selected in this way. But if a skater was selected for more than one distance, the remaining places were filled with qualified skaters who finished third or fourth. The order in which these skaters were selected was decided by the KNSB.

The main problem of this selection method is the fact that all best and second best skaters on a distance were selected automatically, whereas the medal chances of a third best skater on another distance could have been much higher. As the main goal of the KNSB has always been winning as much medals as possible, preferable gold ones, this procedure needed to be changed. Obviously, an inter-distance comparison then becomes necessary.

This chapter presents a new selection procedure that includes win probabilities, with which results on different distances become comparable. The win probabilities are used by the KNSB to select maximal ten male and ten female skaters for the Olympic Winter Games in 2010. In Section 2 is described how the win probabilities are calculated. In Section 3 a theoretical selection procedure is described. In Section 4 the actual selection procedure is given and in Section 5 the results of this procedure are discussed. Section 6 contains conclusions and recommendations.

## 4.2 Medal win probabilities and the 2010 performance matrix

The main subject of this chapter is to compare results from different distances. In the new selection procedure the comparison is based on the medal win probabilities of individual skaters at the Olympics of 2010. We calculate these probabilities from results of international tournaments organized in the year before the Olympics, namely the World Single Distance Championships (WSDCh) of 2009, and the five World Cups Competition (WCC) races between November 2009 and December 2009. For the Dutch skaters we also include the National Championship in Oktober 2009.

Although the data set is small, we have a maximum of six results per skater per distance; using older data will bias the estimates too much, because form and progression of skaters change rather quickly over time. For example, Erben Wennemars realized a world record on the 1500m in November 2008, but one year later he was degraded to the B-group of the World Cup.

### 4.2.1 Mutual differences

The estimation of win probabilities is based on actual skating results. However, in order to use results from different tournaments, we have to correct for the influence of the location of the rink. Results on high altitude rinks are usually much better, than those on low altitude rinks. Moreover, the quality of the ice may vary significantly amongst rinks; see Kamst (2010). Therefore, instead of using actual skating times, we use 'mutual differences', explained below, to remove rink influences.

Define

$$\begin{aligned} K &= \text{set of tournaments;} \\ D_k &= \text{set of distances skated during} \\ &\quad\ \text{tournament } k; \\ I_{dk} &= \text{set of skaters competing on} \\ &\quad\ \text{distance } d \text{ during tournament } k. \end{aligned}$$

For each $k \in K, d \in D_k$, $i \in I_{dk}$, and $a = 1, 2, 3, \ldots$, we also define

$$\begin{aligned} T_{idk} &= \text{result of skater } i \text{ at distance } d \\ &\quad\ \text{during tournament } k; \\ \tilde{T}_{adk} &= \text{a-th fastest result on distance } d \\ &\quad\ \text{during tournament } k; \end{aligned}$$

and

$$AV^5 \quad = \quad T_{idk} - \frac{1}{5}\sum_{a=1}^{5}\tilde{T}_{adk},$$

i.e., $AV^5_{idk}$ is the difference between the result of skater $i$ and the average result of the five best skaters on distance $d$ during tournament $k$. In this chapter the $AV5$-values are used and in Section 2.6.2 we explain why the best five are chosen.

Figure 4.1 shows the $AV5$-values of the 1500m of the six tournaments. The histograms of the other distances for both men and women are shown in Figure 4.2. The shape of most histograms is 'right tailed', indicating that results of top ranked skaters are more dense than those of the sub toppers. It also indicates that it is hard to skate a little faster when being already faster than average. This fact agrees with Gould's hypothesis (see Gould (1999)): due to a virtual boundary it becomes more and more difficult for top athletes to improve, while there occurs congestion in front of this boundary. In Section 2.3.4 we elaborate the relation between the development of skating times and Gould's hypothesis in more detail. The histograms of the longer distances, like the men's 10000m and women's 5000m, due to the low number of observations no a clear right tailed shape. However when we include several more seasons, also these distances show a right tailed shape.

Figure 4.1 also shows five resent results of Stefan Groothuis, a Dutch 1500m specialist. He finishes 3rd, 4th, 5th, 7th, and 21th during the tournaments mentioned above. The $AV5$-values are also shown in Table 4.2. In Table 4.1 it is shown how his $AV5$-value on the second World Cup in Heerenveen (November 9, 2010) is calculated.

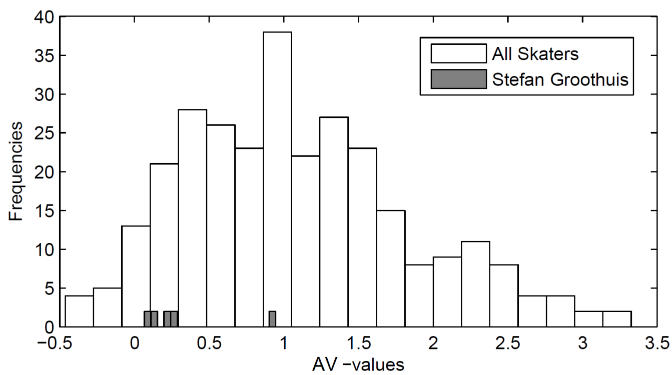



**Figure 4.1.** Histogram 1500m AV5-values

**Table 4.1.** The calculation of AV5-values, Heerenveen, November 9, 2010

| Rank | Skater | Country | Time | r-time | $AV5$-value |
|---|---|---|---|---|---|
| 1 | Shani Davis | USA | 1:44.48 | 34.827 | -0.353 |
| 2 | Havard Bokko | NOR | 1:45.57 | 35.190 | 0.010 |
| 3 | Stefan Groothuis | NED | 1:45.74 | 35.247 | 0.067 |
| 4 | Mark Tuitert | NED | 1:45.83 | 35.277 | 0.097 |
| 5 | Rhian Ket | NED | 1:46.08 | 35.360 | 0.180 |
| Average of best five | | | | 35.180 | |

**Table 4.2.** The 1500m AV5-values of Groothuis

| Date | Tourn. | Position | AV5-value |
|---|---|---|---|
| 11.13.2009 | WC | 3 | 0.07 |
| 11.08.2009 | WC | 4 | 0.12 |
| 12.04.2009 | WC | 5 | 0.23 |
| 03.12.2009 | WSDCh | 7 | 0.27 |
| 11.21.2009 | WC | 21 | 0.95 |

**Individual distributions of AV5-values**

In order to simulate skating times we need to estimate the distribution of the $AV5$-values for all skaters. Based on the shape of the histograms (see Figures 4.2), all showing a right-tailed distribution, we assume that the $AV5$-values follow a log-normal distribution. However, we then need the fact that all values are positive. Since some of the $AV5$-values are negative, we use $AV^5 + M$ instead of $AV^5$, where $M = -\lfloor min(AV^5_{idk}) \rfloor$. Since, $M = 1$, we assume that the values of $(AV^5+1)$ are log-normal distributed. In fact, we assume that for each skater $i$ and each distance $d$, the $(AV^5 + 1)$-values follow a log-normal distribution with mean $\mu_{id}$ and variance $\sigma_{id}$. For example, the 1500m mean and variance values of Groothuis' log-normal distribution are 0.11 and 0.01, respectively, meaning that $\log(AV^5_{Groothuis}+1) \sim N(0.11, 0.01)$. For each distance $d$, the means are estimated by a fixed effect model with the skaters as fixed effects, i.e.,

$$\log(AV^5_{idk} + 1) = \mu_{id} + \epsilon_{idk} \tag{4.1}$$

where $\epsilon$ is a normally distributed error term with variance $\sigma^2_{d\epsilon}$. The estimated variance of skater $i$ on distance $d$ is denoted and defined by

$$\bar{\sigma_{id}}^2 = \frac{\sum_k (\log(AV^5_{idk} + 1) - \hat{\mu_{id}})^2}{m_{id}},$$

where $m_{id}$ is the number of observations for skater $i$ on distance $d$. However, as the number of observations per skater is small, the influence of outliers on the variance may be large. In order to make the estimator of the variance more robust and to reduce the influence of outliers on the variance, we limit a skaters' variance in the following way. The variance of a skater will never be larger than two times the variance of the residuals from regression model 4.1. In case the variance is bigger, we take two times the estimated variance of the population. So for each skater $i$ and distance $d$, we take as estimated variance

$$\hat{\sigma^2_{id}} = min\{\bar{\sigma^2_{id}}, 2\tilde{\sigma^2_{d\epsilon}}\}.$$

In order to increase the influence of results of the more important and more recent tournaments, we use the option of duplicating results, meaning that 'more important' results are included twice in the dataset, and have more influence on the estimations. In our case, the KNSB requested to increase the weights of the WSDCh and the WC in Calgary.

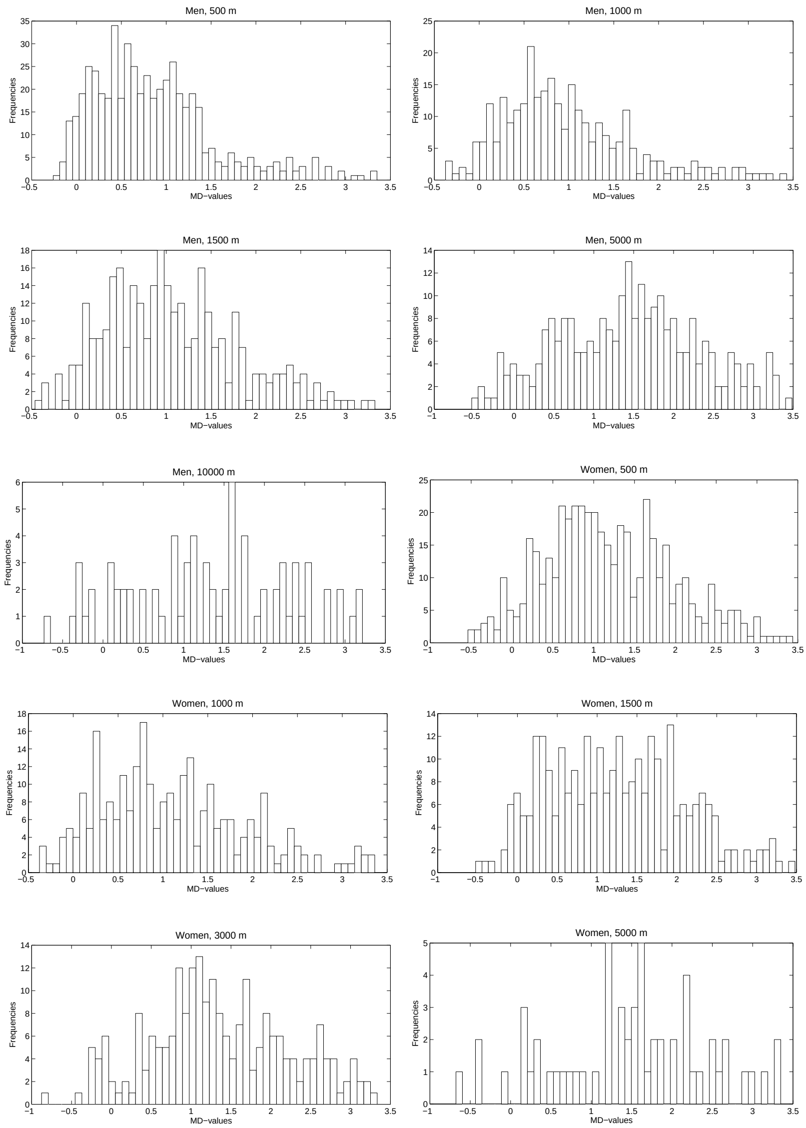



**Figure 4.2.** Histograms $AV5$-values, Men and Women

### 4.2.2 Position probabilities

The estimated log-normal distributions of the skaters are used as input in a simulation model used for generating race results. For each skater, we draw a AV5-value from his distribution, i.e., the simulated AV5-value $\tilde{AV}^5_{id}$ of skater $i$ on distance $d$ is calculated by

$$\tilde{AV}^5_{id} \quad = \quad \exp^{x_{id}} - 1,$$

where $x_{id} \sim logN(\hat{\mu_{id}}, \hat{\sigma_{id}})$. The AV5-values are then ordered into a ranking. This procedure is repeated 5000 times. Then for each distance $d$ we count the number of times skater $i$ has finished on position $m$ during these 5000 race simulations and define this as

$$N_{idm} = \text{number of times skater } i \text{ finished on position } m \text{ at distance } d.$$

These figures are transformed into probabilities by dividing them by the number of simulations, in this case 5000. Define

$$p_{id} \quad = \quad \text{the simulated finish position of skater } i \text{ on distance } d,$$

and

$$PP_{id}^{m} \quad = \quad \text{the probability that skater } i \text{ will finish on position } m \text{ at distance } d.$$

This probability is calculated for each skater $i$ and each distance $d$ by means of

$$PP_{id}^{m} \quad = P(p_{id} = m) = \quad \frac{N_{idn}}{5000}.$$

For example, during the 5000 simulations, Groothuis ended up 391 times on the second place at the 1500m, which means that $PP_{Groothuis,1500m,2} = \frac{391}{5000} = 0.07$.

**Win probabilities**

The win probabilities are the basis of the selection procedure, i.e., the choice of two teams of ten Dutch skaters (men and women) with the highest total win probability. As win probability we take the probability of winning a medal, either gold, silver, or bronze. Hence, it is the probability of finishing within the first three positions of an Olympic distance race. The win probability $PM_{id}^{3}$ of skater $i$ on distance $d$ is defined as

$$PM_{id}^{3} = \sum_{m=1}^{3} PP_{id}^{m} = P(p_{id} \leq 3).$$

In general, we define the probability of skater $i$ of finishing within the first $\alpha$ on distance $d$ as

$$PM_{id}^{\alpha} = \sum_{m=1}^{\alpha} PP_{idm} = P(p_{id} \leq \alpha).$$

**Performance matrices**

The main component in the selection procedure is the so-called Performance Matrix (PM). A PM lists for all skaters and all Olympic distances the probability of winning a medal on that distance. Tables 4.3 and 4.4 give the PMs for both men and women, respectively. For example, in Table 4.3, the number 16 for Groothuis on the 1500m means that he has a 16% chance of being within the best three during the Olympics on the 1500m. This 16% is the sum of the probabilities that he finishes either first, second, or third, namely 0.02%, 7.8%, 8.4%, respectively. In case a skater has a value of zero, it means he can be selected (has a nomination; see Section 4.4.2) for the distance but has no chance of finishing within the best three. Empty cells refer to distances on which the skater did not participate or has no nomination for.

In the PM for women of Table 4.4 we give both the probabilities for being within the best three ($PM^3$) as being within the best six ($PM^6$). This is done because only seven women have a nonzero chance to skate within the best three. Therefore we expand the matrix and include the probabilities of the fourth, fifth and sixth place without a weight. In the sensitivity section, see Section 4.3.3, we discuss several different options to include these probabilities and different weights, and how they affect the outcome.

**Table 4.3.** Performance matrix, $PM^3$, December 2010, Dutch male skaters (in %)

| | **500m** | **1000m** | **1500m** | **5000m** | **10000m** |
|---|---|---|---|---|---|
| Bob de Jong | | | | 71 | 94 |
| Ben Jongejan | | | | 0 | 0 |
| Bob Vries | | | | 4 | 0 |
| Remco Olde Heuvel | | 0 | 4 | | |
| Koen Verweij | | | | 0 | 0 |
| Simon Kuipers | 0 | 17 | 0 | | |
| Stefan Groothuis | 0 | 77 | 16 | | |
| Lars Elgersma | | 1 | 0 | | |
| Jan Blokhuijsen | | | | 16 | 0 |
| Carl Verheijen | | | | 10 | 13 |
| Erben Wennemars | 0 | 0 | 0 | | |
| Rhian Ket | | 0 | 10 | | |
| Wouter Olde Heuvel | | | | 0 | 0 |
| Jacques Koning | 0 | 0 | | | |
| Jan Bos | 0 | 2 | 0 | | |
| Jan Smeekens | 10 | 0 | | | |
| Mark Tuitert | 0 | 9 | 17 | | |
| Sven Kramer | | | 0 | 96 | 100 |
| Ronald Mulder | 5 | 0 | | | |
| Arjen Kieft | | | | | 0 |

**Table 4.4.** Performance matrix, $PM^3$ ($PM^6$), December 2010, Dutch female skaters (in %)

| | 500m | 1000m | 1500m | 3000m | 5000m |
|---|---|---|---|---|---|
| Annette Gerritsen | 10 (83) | 69 (93) | 9 (22) | | |
| Natasja Bruintjes | 0 (0) | 4 (17) | 0 (1) | | |
| Diane Valkenburg | | | 5 (27) | 0 (1) | |
| Anice Das | 0 (0) | | | | |
| Elma Vries | | | 11 (22) | 0 (4) | 0 (8) |
| Gretha Smit | | | | 0 (0) | 0 (0) |
| Ireen Wust | 0 (0) | 1 (6) | 29 (48) | 1 (27) | |
| Jorien Voorhuis | | 0 (0) | 0 (0) | | 0 (1) |
| Laurine Riessen | 0 (1) | 2 (10) | 0 (1) | | |
| Lisette Geest | | | | | |
| Margot Boer | 1 (74) | 8 (64) | 16 (4) | | |
| Marianne Timmer | 1 (19) | 33 (65) | | | |
| Marrit Leenstra | | 0 (0) | 0 (0) | | |
| Moniek Kleinsman | | | | 0 (1) | |
| Paulien Deutekom | | | 0 (0) | | |
| Renate Groenewold | | | | 0 (0) | 0 (14) |
| Sanne Star | 0 (0) | | | | |
| Thijsje Oenema | 0 (1) | 0 (0) | | | |

## 4.3 Selection model

The win probabilities from PMs can be used for the calculation of optimal teams with as objective the maximization of the total win probability. Basically the problem is to assign skaters to positions. The available positions differ among countries and are determined by rules and rankings formulated by the ISU; see Section 4.3.1. For the Olympic Winter Games of 2010, the Netherlands had thirty-five starting tickets, eighteen for the men and seventeen for the women. As is said already, the male skaters have four positions on the 500m, 1000m and the 1500m, plus three on the 5000m and 10000m. The same holds for the women with the exception that they skate the 3000m instead of the 10000m and have only two starting tickets for the 5000m. Furthermore, the ISU has limited the total number of skaters per country to twenty, ten men and ten women, such as to control the total number of skaters at the Olympics; see Section 4.3.1.

The selection problem is closely related to the so called Uncapacitated Facility Location Problem (UFLP) (see, e.g., Galvão and Raggi (1989)): instead of locations we have skaters, while the customers are the starting tickets on the distances. In Sierksma and Boon (2003) a similar method is used to find optimal soccer and volleyball teams. In the human research management literature such models are frequently used to build teams in which the competencies of team members are matched to the tasks of a project; see, e.g., Hlaoittinun *et al.* (2007).

For our Olympic selection problem, PMs are used as incidence matrices between skaters and distances. The complete model is formulated as follows.

**Variables**

For $i \in I, j \in D$, and $\alpha > 1$, the following variables are defined:

$$x_{id} = \begin{cases} 1 & \text{if skater } i \text{ starts at distance } d \\ 0 & \text{otherwise} \end{cases}$$

$$y_i = \begin{cases} 1 & \text{if skater } i \text{ is selected} \\ 0 & \text{otherwise} \end{cases}$$

**Parameters**

$$PM_{id}^{\alpha} = \text{probability that skater } i$$
$$\text{finishes within the best } \alpha \text{ skaters at distance } d.$$

Recall that $PM_{id}^{\alpha}$ is the ($id$)-th entry of $PM^{\alpha}$.

**Objective**

The objective of the selection model reads as follow:

$$Z_{OPT}^{\alpha} = \max_{x_{id}} \sum_{i \in I} \sum_{j \in D} PM_{id}^{\alpha} x_{id}, \tag{4.2}$$

where, in case of maximizing the medal win probability, so $\alpha = 3$, $Z_{OPT}^{3}$ is the maximum total medal win probability.

**Include nominations** In PMs only nominated skaters obtain values, while non-nominated skaters have empty cells. In formula (4.2) however, empty cells are seen as zeros, which means that the model may select non-nominated skaters. For example, the 500m has only two skaters with a positive probability, so it may happen that Sven Kramer, who has no nomination for the 500m, is chosen for the 500m without making the solution infeasible. This is of course an undesirable situation. The exclusion of skaters for their non-nominated distances is modeled as follows.

For $i \in I, j \in D$, and $\alpha > 1$, define

$$c_{id}^{\alpha} = \begin{cases} PM_{id}^{\alpha} & \text{if skater } i \text{ is qualified for distance } d \\ -1 & \text{otherwise} \end{cases}$$

This changes the objective function of (4.2) into

$$Z_{OPT}^{\alpha} = \max_{x_{id}} \sum_{i \in I} \sum_{j \in D} c_{id}^{\alpha} x_{id} \tag{4.3}$$

Now only nominated skaters are selected, since non-nominated skaters decrease the value of the objective function.

**Constraints**

The total win probability is maximized under two constraints, namely the maximum number of participants per distance and the size of the selection. Now let,

$$k_d = \text{maximum number of participants on distance } d,$$

and

$$h = \text{size of the selection team.}$$

The constraints now read as follows:

$$\sum_{i \in I} y_i \leq h \tag{4.4}$$

$$\sum_{i \in I} x_{id} \leq k_d \quad \text{for } d \in D. \tag{4.5}$$

We have already mentioned that $h = 10$ for both men and women; see also Section 4.3.1. Furthermore, to ensure that each skater that is selected for a distance is also selected for the team, and that each skater that is part of the team also skates at least one distance, the following two logical constraints are needed.

$$x_{id} \leq y_i \qquad \text{for } i \in I, d \in D \tag{4.6}$$

$$\sum_{d \in D} x_{id} \geq y_i \qquad \text{for } i \in I. \tag{4.7}$$

Constraint (4.6) ensures that if skater $i$ is selected for a distance ($x_{id} = 1$ for some $d$), he is also selected for the team ($y_i = 1$). Constraint (4.7) ensures that if skater $i$ is selected ($y_i = 1$), he skates at least one distance ($x_{id} = 1$ for some $d$). Note that skater $i$ can start at more than one distance, so that it may happen that $\sum_{d \in D} x_{id} \geq 2$.

### 4.3.1 Number of participants per country

Restrictions, such as the limitation of the number of skaters for men and women per country, are formulated by the ISU in the ”Qualification System for the XXI Winter Olympics Games, Vancouver 2010”; see ISU (2008). In this document the following rules and quotas can be found

- The total number of skaters to be entered into the 2010 Olympic Winter Games Speed Skating events (IOC Quota) is 180;
- The total overall maximum number of entered speed skaters per national ISU member (NF), respectively National Olympic Committee (NOC) (NF/NOC overall quota) are at most twenty skaters per NF/NOC;

- A maximum of ten for women and ten for men for NFs/NOCs, that have been allocated quota places for all events/distances, including the team pursuit events. Other NFs/NOCs have a maximum of eight for both men and women;

- The number of skaters entered per NF/NOC per event may not exceed:
  - four in each of the individual events/distances 500m, 1000m, and 1500 m;
  - three in each of the individual events/distances 3000m (women), 5000m (women and men), and 10000m (men).

- Maximum number of skaters per event/distance (event/distance quota):
  - 500m, 1000m, and 1500m: 36 women and 40 men;
  - 3000m for women and 5000m for men: 28;
  - 5000m for women and 10000m for Men: 16.

The actual so-called quotas (number of participants) per country are a result of the Special Olympic Qualification Classification (SOQC). This SOQC is a combination of the WC rankings of season 2008/2009, and the ranking based on the best times during that WC season. It determines for each country the number of skaters that can start per distance. Details can be found in ISU (2008). Furthermore, the ISU requires that skaters are only allowed to participate in the Olympic Games if they satisfy the so called ISU Qualifying Times in the period between July 1, 2009 through January 17, 2010. The qualification times are given in Table 4.5. For the longest distances, the 10000m for men and the 5000 for women, skaters can qualify on times on both the 5000m and 100000m for men and 3000m and 5000m for women.

**Table 4.5.** ISU Qualifying Times for the 2010 Olympic Winter Games

| Distance | Women | Men |
|---|---:|---:|
| 500m | 39.50 | 36.00 |
| 1000m | 1:18.50 | 1:11.00 |
| 1500m | 2:00.00 | 1:49.00 |
| 3000m | 4:15.00 | n.a. |
| 5000m | 7:20.00 or 4:10.00 (3000m) | 6:35.00 |
| 10000m | n.a. | 13:30.00 or 6:30.00 (5000m) |

The Netherlands is allowed to delegate ten men and ten women in total. They also are allowed to send the maximum number of skaters to each distance with the exception of the women's 5000m. Due to bad results, namely a 11th, 19th, and 24th position during the Hamar WC on November 22, 2009, the Netherlands has only two starting tickets on the women's 5000m.

**The model**

The complete optimization model now reads as follows:

Model OS

$$Z_{OPT}^{\alpha} = \max_{x_{id}} \sum_{i \in I} \sum_{j \in D} c_{id}^{\alpha} x_{id}$$

subject to

$$\begin{aligned}
&\sum_{d \in D} x_{id} \geq y_i \text{ for } i \in I \\
&\sum_{i \in I} x_{id} \leq k_d \text{ for } d \in D \\
&\sum_{i \in I} y_i \leq h \\
&x_{id} \leq y_i \text{ for } i \in I, d \in D \\
&x_{id} \in \{0,1\} \text{ for } i \in I, d \in D \\
&y_i \in \{0,1\} \text{ for } i \in I.
\end{aligned}$$

Using the PMs from Tables 4.3 and 4.4, optimal selections can be calculated with Model OS; see Section 4.3.2.

**Model complexity**

As said, model OS can be seen as an UFLP problem, covering the simple plant location problem, the p-medium problem, and the fixed charge p-medium problem. All these problems are known to be NP-hard; see Galvão and Raggi (1989). In our case, the model has a total of $|I||D|+|I|$ binary decision variables, and $|I|+|D|+|I||D|+1$ constrains. Since $|D| = 5$ and $|I| < 30$, the problem is relatively small in comparison to usual UFLP problems with up to 200 customers and 200 potential facility sites. In Galvão and Raggi (1989) a 3-stage method is presented that solves such large problems within acceptable time limits. Model OS can be solved easy and fast: the results in Section 4.3.2 are produced in less than a second; we used EXCEL and the build-in EXCEL solver for solving the problem.

### 4.3.2 Results

The theoretical optimal selections for both men and women are based on the results of the five WCs in November and December of 2009, the Dutch National Single Distance Championships of 2010, and the WCSD of 2009.

**Men's selection**

The solution for the men's model is given in Table 4.6. This table shows that Ste-

**Table 4.6.** Optimal Selection Men (using $PM^3$)

<table>
<tr><th>500m</th><th>1000m</th><th>1500m</th><th>5000m</th><th>10000m</th></tr>
<tr><td colspan="3">Simon Kuipers</td><td colspan="2">Bob de Jong</td></tr>
<tr><td colspan="3">Stefan Groothuis</td><td colspan="2">Sven Kramer</td></tr>
<tr><td>Jan Smeekens</td><td colspan="2">Mark Tuitert</td><td>Jan Blokhuizen</td><td>Carl Verheijen</td></tr>
<tr><td>Ronald Mulder</td><td colspan="2">Rhian Ket</td><td></td><td></td></tr>
</table>

fan Groothuis, Rhian Ket and Simon Kuipers are, besides other distances, selected for the 500m, 1000m, and 1500m, respectively. However, none of these skaters has a chance of winning a medal on that distance; they are chosen arbitrarily from the nominated skaters with a win probability of zero. For example, either Jan Smeekens or Ronald Mulder, who both have a nomination on the 1000m, can replace Ket on the 1000m without changing the optimal value.
If we want a unique optimal solution without the arbitrary choice between two or more nominated skaters with $c^3_{id} = 0$, we need to extend the model by including probabilities of finishing on the fourth, fifth, or sixth position. In order to keep the original objective of winning as many medals as possible, these probabilities need to be weighted much less in the objective function than the original medal win probabilities.

For each $i \in I$, and $d \in D$,

$$c^{3,6}_{id} = \begin{cases} PM^3_{id} + 0.01(PM^6_{id} - PM^3_{id}) & \text{if skater } i \text{ is nominated for distance } d \\ -1 & \text{otherwise} \end{cases}$$

So, for nominated skaters, $c^{3,6}_{id}$ is the probability that skater $i$ wins a medal on distance $d$ plus a small portion of the probability that he finishes on the fourth, fifth, or sixth place. Again, $c^{3,6}_{id} = -1$ if skater $i$ is not nominated for distance $d$.

The results of this adjustment are shown in Table 4.7. The solution shows that

**Table 4.7.** Optimal Selection Men (using $PM^{3,6}$, weighted)

<table>
<tr><th>500m</th><th>1000m</th><th>1500m</th><th>5000m</th><th>10000m</th></tr>
<tr><td colspan="2">Simon Kuipers</td><td colspan="3">Sven Kramer</td></tr>
<tr><td colspan="3">Stefan Groothuis</td><td colspan="2">Bob de Jong</td></tr>
<tr><td>Jan Smeekens</td><td colspan="2">Mark Tuitert</td><td>Jan Blokhuizen</td><td>Carl Verheijen</td></tr>
<tr><td colspan="2">Ronald Mulder</td><td>Rhian Ket</td><td></td><td></td></tr>
</table>

Kuipers is replaced by Kramer on the 1500m. Kramer has a slightly better chance to finish within the best six, namely 1.0% against 0.0%. The fourth position on the 1000m is also changed: Mulder is selected at the cost of Ket. The selection of Groothuis on the 500m remains unchanged.

**Women's selection**

In Table 4.8 the optimal selection of the women based on $PM^6$ is given. This selection is the same as the one for $PM^3$ with weighted positions four through six with a weight of 0.01. Although, Marianne Timmer was injured during the December OQT, we have not excluded her from the dataset, mainly because the KNSB has given her a second chance during the January 2010 OQT II; see Section 4.4.5. In case we exclude Timmer, she is replaced on both the 500m and the 1000m by Laurine van Riesen.

Table 4.9 shows the results when three starting positions on the 5000m would have been available. It turns out that Jorien Voorhuis then obtains the third position on the 5000m at the cost of Moniek Kleinsman's 3000m. Voorhuis has a 0.9% chance to finish within the best six on the 5000m, whereas Kleinsma has only a 0.5% chance on the 3000m. On the 3000m, Kleinsman is replaced by Diane Valkenburg, who already skates the 1500m.

**Table 4.8.** Optimal Selection Women (using $PM^6$)

<table>
<tr><th>500m</th><th>1000m</th><th>1500m</th><th>3000m</th><th>5000m</th></tr>
<tr><td>Thijsje Oenema</td><td>Natasja Bruintjes</td><td>Diane Valkenburg</td><td>Moniek Kleinsma</td><td>Renate Groenewold</td></tr>
<tr><td colspan="2">Marianne Timmer</td><td colspan="3">Elma de Vries</td></tr>
<tr><td colspan="2">Margot Boer</td><td colspan="2">Ireen Wüst</td><td></td></tr>
<tr><td colspan="3">Annette Gerritsen</td><td></td><td></td></tr>
</table>

**Table 4.9.** Optimal Selection Women (using $PM^6$, three 5000m positions)

<table>
<tr><th>500m</th><th>1000m</th><th>1500m</th><th>3000m</th><th>5000m</th></tr>
<tr><td colspan="2">Margot Boer</td><td colspan="3">Elma de Vries</td></tr>
<tr><td colspan="2">Marianne Timmer</td><td colspan="2">Ireen Wüst</td><td>Renate Groenewold</td></tr>
<tr><td>Thijsje Oenema</td><td>Natasja Bruintjes</td><td colspan="2">Diane Valkenburg</td><td>Jorien Voorhuis</td></tr>
<tr><td colspan="3">Annette Gerritsen</td><td></td><td></td></tr>
</table>

### 4.3.3 Various objective functions

In Section 4.3.2 we optimized the total win probability for each medal having the same weight, i.e., winning a gold medal has the same weight as winning a silver, and bronze medal. However, one may argue that gold medals are preferred over silver or bronze medals. On the other hand, unlike the Netherlands, where it is priority one to win speed skating medals, other countries may be satisfied with a position within the best six. These options can be included in the objective function by weighing the probabilities of the medal positions in the following way. Let $w_m$ be the weight for position $m$. Then the weighted value $c_{id}^{3,6}$ can be defined as

$$c_{id}^{3,6} = \begin{cases} \sum_{m=1}^{6} w_m PP_{id}^m & \text{if skater } i \text{ is nominated for distance } d \\ -1 & \text{otherwise} \end{cases}$$

If one values gold medals much higher than silver and bronze ones, the weight of winning a gold medal should exceed the weights of silver and bronze medals. In Table 4.10 we have listed three weight schedules, namely:

- Alternative 1: Linear medal weights;
- Alternative 2: Exponential medal weights;
- Alternative 3: Equal weights for the first six positions.

**Table 4.10.** Weight values, $w_p$

| | **Gold** $p = 1$ | **Silver** $p = 2$ | **Bronze** $p = 3$ | **4th, 5th, 6th.** $p = 4, 5, 6$ |
|---|---|---|---|---|
| Model OS | 1 | 1 | 1 | 0.01 |
| Alternative 1 | 3 | 2 | 1 | 0.01 |
| Alternative 2 | 100 | 10 | 1 | 0.01 |
| Alternative 3 | 1 | 1 | 1 | 1 |

Alternative 1 uses a linear weight system for the three medals, which means that the difference in weight between the medals is a constant. In Alternative 2, we use an exponential weight system in which the weight increases exponentially with the value of the medal. In Alternative 3, we use $PM^6$ in the objective function and maximize the probability of finishing within the best six.

In case of Alternative 1 and 2, the optimal solution of Table 4.7 does not change. All skaters with a high gold win probability are already selected and a change in weighing the medals does not result in the selection of other skaters. In case of Alternative 3, there are just a few changes; see Table 4.11. By including the fourth,

**Table 4.11.** Optimal Selection, Alternative 3

<table>
<tr><th>500m</th><th>1000m</th><th>1500m</th><th>5000m</th><th>10000m</th></tr>
<tr><td colspan="2">Simon Kuipers</td><td>Rhian Ket</td><td colspan="2">Bob de Jong</td></tr>
<tr><td colspan="3">Stefan Groothuis</td><td colspan="2">Sven Kramer</td></tr>
<tr><td>Jan Smeekens</td><td colspan="2">Mark Tuitert</td><td colspan="2">Carl Verheijen</td></tr>
<tr><td>Ronald Mulder</td><td colspan="2">Remco Olde Heuvel</td><td></td><td></td></tr>
</table>

fifth and sixth positions with equal weights the model selects the more 'stable' performing skaters. Carl Verheijen has a fair chance of finishing within the best six on both the 5000m (52%) and the 10000m (84%). He replaces Jan Blokhuizen, who only has results on the 5000m and a chance of 54% of being within the best six. Since Verheijen is already selected for the 10000m in the original situation, another tenth skater can be selected to increase the total win probabilty. The new selected skater is Remco Olde Heuvel, who replaces Mulder on the 1000m and Kramer on the 1500m. Where Mulder has a 0% chance and Kramer has only a 1% chance to finish within the best six, Olde Heuvel has a 0.1% chance on the 1000m, and a 25% chance on the 1500m. For the women, the original and Alternative 3 give the same solution. The same holds for Alternatives 1 and 2.

## 4.4 Selection procedure in practice

For the KNSB clear rules, equal opportunities, and a fixed performance moment, namely the Olympic Qualification Tournament (OQT), are the key drivers for the final selection. Already two years before the Olympics the KNSB has specified the tournaments that are important for the final selection, and how nominated skaters should perform there for getting a so called qualification. Such a qualification is very important for the final selection during the OQT. The performances during these important tournaments are used in the PMs. Based on these PMs, a so called Performance Ranking is made. The clear rules enable trainers and skaters to make training schedules that lead to performance peaks at the right times. So with the December OQT as a focus point and the Performance Ranking as the main selection criterion, a clear and fair route is created for skaters to make it to the Olympics.

### 4.4.1 Stakeholders in the procedure

The KNSB is not the only policy maker concerning the selection of skaters for Olympic Winter Games. Also the IOC (International Olympic Committee) and the ISU (International Speed Skating Union) have a say in the selection. These organizations formulate nomination criteria, namely for each skating distance they set a fixed time limit that needs to be reached by a skater in order to obtain a nomination for that distance. Only skaters that are nominated can be selected by national organizations. Of course, this is to exclude skaters with very poor performances. For the Netherlands, the main decision makers are the NOC*NSF (Dutch Olympic Committee/Dutch Sport Federation) and the KNSB. The actual selection procedure can be formulated as a sequence of decision steps. Below we present and describe these steps; see also KNSB (2009).

### 4.4.2 2010 KNSB selection procedure

The KNSB wants to take into account the fact that skaters may have their peak just before the Games. So they want to take the final decisions as late as possible. This is why they organize the OQT in December. Almost all selection decisions are taken based on the ranking results of this OQT.

**The 2010 selection procedure**

**Step 1. IOC-ISU Nomination** Special time limits are set for each distance by the IOC and the ISU. In order to receive a so called nomination for a distance at least the time limit for that distance needs to be skated in the period July 1, 2009 through January 17, 2010, during an official ISU tournament (ISU Regulations: communication No. 1572). Only nominated skaters can be selected for the Games by the national organizations. The total number of skaters that a country is allowed to delegate is determined by the Special Olympic Qualification Classifications (see Section 4.3.1). Recall that for the Netherlands the current amounts are ten for the men and ten for the women, being the maximum numbers any country is allowed to select.

**Step 2. Qualification** As the Netherlands has many talented skaters (almost all are professionals), also the number of nominated skaters is high. At the end of December 2009 at least 33 men and 34 women were nominated. Therefore the KNSB needs to use its own system to select skaters from the nomination list.

**Step 2a. Pre-qualification** A first pre-selection of the nominated skaters is made based on results obtained between 2008 and 2010. Each pre-selected skater receives the status of qualified skater. This pre-qualification can be obtained by satisfying at least one of the following criteria.

- A top-3 classification during the WSDCh 2009;
- A top-8 classification during one of the five WCs, organized between November 2009 and December 2009;
- A top-8 classification during either one of the WCs, or during the WSDCh in the season 2008/2009, in both cases combined with a top-12 classification during one of the WCs in the season 2009/2010 on the same distance;
- Winning a distance during the OQT in December 2009.

**Step 2b. Special status** A skater who became world champion during the WSDCh 2009 and wins one of the five WCs in the Olympic season 2009/2010 obtains a special status. This means that the KNSB has the possibility to select this skater immediately for the Games, without letting him/her compete in the December OQT; see Step 4a.

**Step 3a. PM ranking** [10] While a PM consists of numbers reflecting the win probabilities of skaters, the PM ranking is an ordered list of the 18 starting positions for the men, and of the 17 starting positions for the women. Recall (see Section 4.3.1) that the women have lost one starting position on the 5000m. The PM rankings are designed in the following way. The first position in a PM ranking is the distance that corresponds to the highest win probability in the corresponding PM. Then the

[10] In the Dutch media a PM ranking was referred to as a Performance Matrix (in Dutch: Prestatiematrix)

distance with the second highest win probability is taken and put on the second position. This procedure is continued until the ranking has reached the maximum number of starting positions on that distance (for example, this maximum is four for the 1500m; see Section 4.3.1).
**Step 3b. PM-adjustments** The KNSB has the right to change the ordering of the PM rankings, for instance based on special arguments of coaches. So, PMs serve as decision support tools: the KNSB stays responsible for the final PM ranking.

**Step 4. The Olympic Qualification Tournament (OQT)** The final PM ranking is filled with names of skaters. The two resulting lists of 18 male skaters and 17 female skaters are called *Selection Rankings*. The assignment of skaters to a PM ranking is carried out in the following way.
**a. Special status** Special status skaters (see Step 2b) are assigned first. They take the first positions of the distance on which they have a special status.
The remaining places on the PM rankings are filled in by using the final results of the December OQT. This is done by distinguishing the following two situations.
**b1. Qualified skater wins OQT distance** If a qualified skater wins an OQT distance and there is an open position on the PM ranking for this distance after Step 4a, then this skater is assigned to that distance position. If there is still a position available on that distance, the next qualified skater from the OQT result of that distance is taken. (So possible higher ranked non-qualifiers are neglected.). This procedure is repeated until all positions of the PM ranking for that distance are filled. In case there are not enough qualified skaters, the highest ranked non-qualifiers are filling the remaining positions.
**b2. Non-qualified skater wins OQT distance** If a non-qualified skater wins a distance at this OQT, he/she receives the status of 'newly qualified skater'. If this happens, then the KNSB keeps the last position at that distance open, and has to organize a second OQT in Januari 2010; see Step 5. The other places are filled with the highest ranked qualified skaters.

**Step 5. Second OQT** The KNSB organizes a second OQT in January 2009 just before the Games when either one of the following situations occurs.
**a.** A qualified skater is sick or injured and cannot participate in the first OQT;
**b.** A non-qualified skater was the winner of an OQT distance; see Step 4(b2).
For each distance with either a newly qualified skater or one that was sick or injured during the first OQT, a competition is organized together with all other qualified skaters that are not yet selected. If a sick or injured qualified skater turns out be the only participant, he automatically occupies the last open position. Otherwise, winners of the second OQT fill the positions that are kept open in Step 4.

**Step 6. More than ten skaters are selected** After applying the previous steps it may happen that more than ten skaters are selected. In this case, the eleventh and higher selected skaters are removed from the PM list. They are replaced by skaters who are already selected within the first ten on other distances. First, the highest ranked qualified skaters during the OQT are selected and then the non-qualified skaters.

**Step 7. Team pursuit** (see Section 4.4.3) During Step 6, the KNSB can choose to select only nine skaters and leave the last position open for a non-selected skater who will start in the team pursuit. This skater will be assigned to an open position on a certain distance and is especially selected to increase the win probability of the team pursuit. Due to the restriction of ten skaters per country, this would mean that the skater selected for the team pursuit replaces the tenth skater of the PM ranking.

### 4.4.3 Team pursuit

Besides the five individual distances, there is another skating event at the Olympics, namely the team pursuit. In the team pursuit a team of three skaters compete against another in a knock-out tournament. Each race exists of eight (six for women) laps, where the teams start on the opposite straight lanes of the track. The winner of the race advances to the next round. The ISU has decided that members of a team pursuit have to participate on one of the individual distances. The KNSB has decided to build a team from the selected skaters on the 1500m and the 3000m/5000m, and to neglect the team pursuit in the selection process. As explained in Step 7, only in case the pursuit team does not satisfy the desired quality level, the KNSB may decide to select a special skater for the team pursuit and let him/her start at one of the individual distances, at the cost of a skater who was already qualified.

### 4.4.4 2010 PM rankings: KNSB and 'Optimal'

The rankings are calculated for two situations, namely for all tournaments through the December 2009 WC of Calgary, and all tournaments through the December 2009 WC of Salt Lake City. Table 4.12 shows the final KNSB PM ranking (columns two and five) and the probabilities until both WCs. The KNSB has based its final rankings mainly on the results until Calgary, because many Dutch skaters were not present in Salt Lake City. During the Salt Lake City WC these skaters were already in training for the important December OQT. In comparison with the calculated 'until Calgary', the KNSB has made the following adjustments. The women's PM 1500m-nr4 (11.5%) is in the KNSB ranking on the fifteenth position, while in the 'until Calgary' ranking this is position thirteen. The reason for this adjustment is a consequence of a comparison between the original PM with and the PM without the October National Championships (NK). In the PM with the NK, the fourth position on the 1500m has a much higher probability (namely 11.5%) than the same position in the PM without the NK (namely 4.0%). This is caused by the performance of Annette Gerritsen, who won the NK but focused during the WC on the 500m and the 1000m, rather than on the 1500m. All other positions in both women's PMs don't differ much. Remarkable in the men's PM ranking is the fact that the first 500m ticket is not one of the

**Table 4.12.** 2010 PM ranking

| | Men | | | Women | | |
|---|---|---|---|---|---|---|
| | KNSB | until Calgary | until Salt Lake City | KNSB | until Calgary | until Salt Lake City |
| 1 | 10000m-nr1 | 100.0% | 100.0% | 1000m-nr1 | 95.9% | 93.2% |
| 2 | 5000m-nr1 | 97.5% | 96.1% | 500m-nr1 | 81.9% | 83.3% |
| 3 | 1000m-nr1 | 85.9% | 76.6% | 1000m-nr2 | 82.2% | 65.4% |
| 4 | 10000m-nr2 | 93.8% | 94.2% | 1000m-nr3 | 74.2% | 64.2% |
| 5 | 5000m-nr2 | 70.6% | 71.2% | 500m-nr2 | 71.5% | 74.1% |
| 6 | 1000m-nr2 | 21.3% | 16.9% | 1500m-nr1 | 64.3% | 48.9% |
| 7 | 1500m-nr1 | 17.3% | 17.2% | 1500m-nr2 | 37.7% | 27.8% |
| 8 | 1500m-nr2 | 15.9% | 16.3% | 1000m-nr4 | 22.6% | 17.8% |
| 9 | 1000m-nr3 | 13.9% | 8.9% | 500m-nr3 | 18.5% | 19.1% |
| 10 | 10000m-nr3 | 13.6% | 13.0% | 1500m-nr3 | 17.3% | 22.5% |
| 11 | 5000m-nr3 | 13.0% | 15.7% | 3000m-nr1 | 14.9% | 27.8% |
| 12 | 500m-nr1 | 10.6% | 10.3% | 5000m-nr1 | 12.8% | 14.0% |
| 13 | 1500m-nr3 | 10.2% | 9.9% | 3000m-nr2 | 1.6% | 4.9% |
| 14 | 1000m-nr4 | 4.0% | 2.1% | 5000m-nr2 | 8.0% | 8.4% |
| 15 | 500m-nr2 | 6.8% | 5.1% | 1500m-nr4 | 11.5% | 22.1% |
| 16 | 1500m-nr4 | 4.6% | 3.6% | 500m-nr4 | 0.1% | 1.0% |
| 17 | 500m-nr3 | 0.0% | 0.1% | 3000m-nr3 | 0.8% | 0.5% |
| 18 | 500m-nr4 | 0.0% | 0.0% | | | |

"until x" means all tournaments through WC Calgary, 12-4-2009, or WC Salt Lake City, 12-12-2009.

first ten positions: the first 500m is at position twelve. So it could be possible that the best Dutch male 500m skater would not be present at the Olympics. Compared to previous Olympic selections this is a major change, because in all previous editions of the Olympics Winter Games, the Netherlands are represented on each distance by its best skaters. Looking at the win probabilities, it becomes clear that the distances with relatively strong Dutch performance are chosen at the cost of the weaker 500m.

For the women's PM ranking the long distances are in trouble. The first 3000m and the first 5000m occur on the eleventh and twelfth position, respectively.

### 4.4.5 OQT results and final selection

Only Sven Kramer has obtained a special Olympic status (see Section 4.4.2, Step 2b). Therefore, he did not need to compete in the December OQT. Based on his extraordinary results he was directly selected for the 5000m and 10000m. The winners of the 5000m and 10000m December OQT take the positions 5000m-nr2 and 10000m-nr2 in the PM ranking, respectively. The complete list of the results of the December OQT is given in Tables 4.13 and 4.14. At the start of the tournament, eighteen men and ten women received already at least one qualification; in both tables this is labeled as Q.

Table 4.15 contains the final KNSB selection. We explain the contents of Table 4.15 by going through the steps of Section 4.4.2 in case of the men. First all non-qualified skaters that did not won an OQT distance are removed from the results of Table 4.13, and the remaining qualified skaters are placed in the PM ranking accordingly. For example, in case of the 500m, Ronald Mulder is selected as the second 500m skater. Although he finished third on the men's 500m, behind Simon Kuipers, he is taken as

**Table 4.13.** December OQT results, men
(Q=qualified)

500m

| Name | | Time |
|---|---|---|
| Jan Smeekens | Q | 70.190 |
| Simon Kuipers | | 70.280 |
| Ronald Mulder | Q | 70.380 |
| Jan Bos | | 70.490 |
| Stefan Groothuis | | 70.630 |
| Mark Tuitert | | 70.670 |
| Lars Elgersma | | 70.910 |
| Michel Mulder | | 70.980 |
| Michael Poot | | 71.280 |
| Beorn Nijenhuis | | 71.360 |
| Sjoerd de Vries | | 71.740 |
| Kjeld Nuis | | 71.850 |
| Pim Schipper | | 72.120 |
| Hein Otterspeer | | 72.220 |
| Sietse Heslinga | | 72.420 |

1000m

| Name | | Time |
|---|---|---|
| Stefan Groothuis | Q | 1:08.50 |
| Simon Kuipers | Q | 1:08.52 |
| Kjeld Nuis | | 1:09.28 |
| Mark Tuitert | Q | 1:09.32 |
| Beorn Nijenhuis | | 1:09.41 |
| Jan Bos | Q | 1:09.42 |
| Lars Elgersma | Q | 1:09.45 |
| Erben Wennemars | | 1:09.60 |
| Ronald Mulder | | 1:09.67 |
| Rhian Ket | | 1:09.68 |
| Sjoerd de Vries | | 1:09.77 |
| Pim Schipper | | 1:10.07 |
| Michel Mulder | | 1:10.37 |
| Remco Olde Heuvel | Q | 1:10.38 |
| Berden de Vries | | 1:10.56 |
| Jan Smeekens | | 1:10.75 |
| Hein Otterspeer | | 1:11.85 |

1500m

| Name | | Time |
|---|---|---|
| Simon Kuipers | Q | 1:45.39 |
| Stefan Groothuis | Q | 1:45.72 |
| Mark Tuitert | Q | 1:46.47 |
| Wouter Olde Heuvel | | 1:46.59 |
| Sven Kramer | Q | 1:46.66 |
| Erben Wennemars | Q | 1:46.67 |
| Remco Olde Heuvel | Q | 1:46.80 |
| Kjeld Nuis | | 1:46.89 |
| Rhian Ket | Q | 1:47.18 |
| Beorn Nijenhuis | | 1:47.56 |
| Renz Rotteveel | | 1:47.97 |
| Ted-Jan Bloemen | | 1:48.09 |
| Tim Roelofsen | | 1:48.11 |
| Ben Jongejan | | 1:48.24 |
| Carl Verheijen | | 1:48.34 |
| Pim Schipper | | 1:48.63 |

5000m

| Name | | Time |
|---|---|---|
| Bob de Jong | Q | 6:14.12 |
| Carl Verheijen | Q* | 6:15.70 |
| Arjen v/d Kieft | | 6:20.14 |
| Jan Blokhuijsen | Q | 6:22.23 |
| Wouter Olde Heuvel | Q | 6:25.41 |
| Ted-Jan Bloemen | | 6:25.48 |
| Renz Rotteveel | | 6:25.73 |
| Tim Roelofsen | | 6:30.98 |
| Ben Jongejan | | 6:31.61 |
| Boris Kusmirak | | 6:35.86 |
| Tom Schuit | | 6:38.46 |
| Mark Ooijevaar | | 6:38.98 |
| Wierd Osinga | | 6:45.79 |
| | | |
| Bob de Vries | Q | Sick |
| Koen Verweij | Q | Sick |

10000m

| Naam | | Tijd |
|---|---|---|
| Bob de Jong | Q | 12:53.63 |
| Arjen v/d Kieft | Q | 13:02.99 |
| Carl Verheijen | Q | 13:11.01 |
| Ted-Jan Bloemen | | 13:20.22 |
| Mark Ooijevaar | | 13:28.60 |
| Willem Hut | | 13:28.84 |
| Wouter Olde Heuvel | Q | 13:30.41 |
| Ben Jongejan | | 13:36.21 |
| Tom Schuit | | 13:47.28 |

**Table 4.14.** December OQT results, women

500m

| Name | | Time |
|---|---|---|
| Annette Gerritsen | Q | 76,16 |
| Margot Boer | Q | 76,38 |
| Laurine van Riessen | Q | 77,5 |
| Thijsje Oenema | Q | 77,55 |
| Natasja Bruintjes | | 78,14 |
| Ireen Wüst | | 78,62 |
| Sanne van der Star | | 78,88 |
| Sophie Nijman | | 79,04 |
| Lotte van Beek | | 79,06 |
| Frederika Buwalda | | 79,41 |
| Ingeborg Kroon | | 79,55 |
| Anice Das | | 79,57 |
| Roxanne van Hemert | | 79,88 |
| Jorien Kranenborg | | 80,02 |
| Tosca Hilbrands | | 80,24 |
| Janine Smit | | 80,33 |
| Marianne Timmer | Q | Injured |

1000m

| Name | | Time |
|---|---|---|
| Margot Boer | Q | 1:15.70 |
| Laurine van Riessen | Q | 1:15.93 |
| Annette Gerritsen | Q | 1:16.57 |
| Natasja Bruintjes | Q | 1:16.89 |
| Ireen Wüst | Q | 1:16.94 |
| Sophie Nijman | | 1:16.96 |
| Lotte van Beek | | 1:17.29 |
| Ingeborg Kroon | | 1:17.48 |
| Marrit Leenstra | | 1:17.74 |
| Roxanne van Hemert | | 1:17.87 |
| Thijsje Oenema | | 1:18.37 |
| Janine Smit | | 1:18.50 |
| Linda de Vries | | 1:18.62 |
| Paulien van Deutekom | | 1:18.64 |
| Maren van Spronsen | | 1:19.07 |
| Jorien Kranenborg | | 1:19.16 |
| Tosca Hilbrands | | 1:19.61 |
| Marianne Timmer | Q | Injured |

1500m

| Name | | Time |
|---|---|---|
| Ireen Wüst | Q | 1:56.97 |
| Annette Gerritsen | Q | 1:57.58 |
| Laurine van Riessen | Q | 1:58.35 |
| Margot Boer | Q | 1:58.38 |
| Lotte van Beek | | 1:58.40 |
| Diane Valkenburg | Q | 1:58.61 |
| Paulien van Deutekom | | 1:59.65 |
| Roxanne van Hemert | | 1:59.72 |
| Jorien Voorhuis | | 1:59.92 |
| Ingeborg Kroon | | 2:00.34 |
| Marrit Leenstra | | 2:01.30 |
| Linda Bouwens | | 2:01.42 |
| Yvonne Nauta | | 2:01.44 |
| Annouk v/d Weijden | | 2:01.57 |
| Janneke Ensing | | 2:02.34 |
| Elma de Vries | Q | 2:02.39 |

3000m

| Name | | Time |
|---|---|---|
| Ireen Wüst | Q | 4:06.80 |
| Diane Valkenburg | Q | 4:09.14 |
| Yvonne Nauta | | 4:10.72 |
| Elma de Vries | | 4:11.43 |
| Moniek Kleinsman | | 4:11.55 |
| Paulien van Deutekom | | 4:12.02 |
| Jorien Voorhuis | | 4:12.38 |
| Janneke Ensing | | 4:12.63 |
| Renate Groenewold | Q | 4:16.07 |
| Lotte van Beek | | 4:17.00 |
| Linda de Vries | | 4:17.82 |
| Annouk v/d Weijden | | 4:18.27 |
| Irene Schouten | | 4:21.07 |
| Linda Bouwens | | |

5000m

| Name | | Time |
|---|---|---|
| Elma de Vries | Q | 7:07.41 |
| Jorien Voorhuis | | 7:12.39 |
| Diane Valkenburg | | 7:13.97 |
| Janneke Ensing | | 7:17.37 |
| Moniek Kleinsman | | 7:20.14 |
| Mireille Reitsma | | 7:21.10 |
| Maria Sterk | | 7:22.37 |
| Yvonne Nauta | | 7:22.90 |
| Annouk v/d Weijden | | 7:25.48 |
| Linda Bouwens | | 7:35.59 |

the second 500m skater, because he has a qualification, while Kuipers has not. The same holds for Mark Tuitert and Jan Bos on the 1000m, and for Sven Kramer on the 1500m. In case of the 5000m, the KNSB had to apply Step 5. As two qualified skaters were sick during the December OQT, a second OQT needed to be organized. For Carl Verheijen, who finished only one second behind the winner Bob de Jong, the second OQT became a deception: he was beaten by Jan Blokhuijsen, who took the last Olympic ticket.

Also for the women's 500m and 1000m, a second OQT needed to be organized. Due to a severe injury, the already qualified Marianne Timmer, could not participate in the December OQT. Unfortunately, the second OQT came to early for Timmer: she was beaten by Ireen Wüst on the 1000m, and by Thijsje Oenema on the 500m.

Since in both PM rankings no more than ten skaters are used, the result of Table 4.15 is the final selection.

**Table 4.15.** The KNSB final selection

| | Men | | | Women | | |
|---|---|---|---|---|---|---|
| | KNSB PM ranking | Skater | # Sel | KNSB PM ranking | Skater | # Sel |
| 1 | 10000m-nr1 | S. Kramer | 1 | 1000m-nr1 | M. Boer | 1 |
| 2 | 5000m-nr1 | S. Kramer | 1 | 500m-nr1 | A. Gerritsen | 2 |
| 3 | 1000m-nr1 | S. Groothuis | 2 | 1000m-nr2 | L. van Riessen | 3 |
| 4 | 10k-nr2 | B. de Jong | 3 | 1000m-nr3 | A. Gerritsen | 3 |
| 5 | 5k-nr2 | B. de Jong | 3 | 500m-nr2 | M. Boer | 3 |
| 6 | 1000m-nr2 | S. Kuipers | 4 | 1500m-nr1 | I. Wüst | 4 |
| 7 | 1500m-nr1 | S. Kuipers | 4 | 1500m-nr2 | A. Gerritsen | 4 |
| 8 | 1500m-nr2 | S. Groothuis | 4 | 1000m-nr4* | I. Wüst | 4 |
| 9 | 1000m-nr3 | M. Tuitert | 5 | 500m-nr3 | L. van Riessen | 4 |
| 10 | 10000m-nr3 | A. van der Kieft | 6 | 1500m-nr3 | L. van Riessen | 4 |
| 11 | 5000m-nr3* | J. Blokhuijsen | 7 | 3000m-nr1 | I. Wüst | 4 |
| 12 | 500m-nr1 | J. Smeekens | 8 | 5000m-nr1 | E. de Vries | 5 |
| 13 | 1500m-nr3 | M. Tuitert | 8 | 3000m-nr2 | D. Valkenburg | 6 |
| 14 | 1000m-nr4 | J. Bos | 9 | 5000m-nr2 | J. Voorhuis | 7 |
| 15 | 500m-nr2 | R. Mulder | 10 | 1500m-nr4 | M. Boer | 7 |
| 16 | 1500m-nr4 | S. Kramer | 10 | 500m-nr4* | T. Oenema | 8 |
| 17 | 500m-nr3 | S. Kuipers | 10 | 3000m-nr3 | R. Groenewold | 9 |
| 18 | 500m-nr4 | J. Bos | 10 | | | |

x* means filled in after the January OQT. # Sel = the number of selected skaters.

### 4.4.6 Comparing the final and the calculated selections

In this section we compare the calculated PM selections, as presented in Table 4.6, with the actual KNSB selection from Table 4.15. Below we show the points where the KNSB selection deviates from the calculated PM selection.

1. 500m: Jan Bos instead of Stefan Groothuis;
2. 1000m: Jan Bos instead of Rhian Ket;
3. 1500m: Sven Kramer instead of Rhian Ket;
4. 10000m: Arjen vd Kieft instead of Carl Verheijen.

Two of these differences are related to the performance of Rhian Ket, who was the national champion on the 1500m in November 2009, and scored a number of good results during the WCs. These performances resulted in a chance of 10% (see Table 4.3) of winning an Olympic medal. However, it turned out that Ket 'peaked' to early: he finished ninth in the December OQT. Jan Bos, who is not in the calculated theoretical selection, replaced Ket. Clearly Bos focussed on the December OQT, while his results during the previous WCs were much worse. He managed to recieve an Olympic ticket for both the 500m and 1000m.

In case of the women, the deviations from the calculated selection (Table 4.8) with the actual KNSB selection (Table 4.15) are listed below.

1. 500m: Laurine van Riessen instead of Marianne Timmer;
2. 1000m: Laurine van Riessen instead of Marianne Timmer;
3. 1000m: Ireen Wüst instead of Natasja Bruintjes;
4. 1500m: Laurine van Riesen instead of Elma de Vries;
5. 1500m: Margot Boer instead of Diane Valkenburg;
6. 3000m: Renate Groenewold instead of Elma de Vries.

As mentioned already, Marianne Timmer was injured during the December OQT and was given a second chance at the January OQT. However she did not manage to win either the 500m or the 1000m, and lost her places in favor of Laurine van Riessen. In case we remove Marianne Timmer from the data set, then also Laurine van Riessen would have been selected on both distances; see Section 4.3.2.
Elma de Vries is originally a marathon skater and did not participate on the 3000m WCs. Although she beated Renate Groenewold on the 3000m (see Table 4.14), the fact that she had no qualification for this distance made that Groenewold, with her ninth place during the December OQT, was selected instead.

## 4.5 Discussion

The main instrument for deciding on the selection of speed skaters for the 2010 Olympic Winter Games was the PM ranking; a list of ordered distances with the highest win-probabilities at the top. In Section 4.2 it is explained how these win probabilities are calculated. In Section 4.4.6 we have seen that the selection calculated from these win probabilities (without the two OQTs) differs from the final KNSB selection on three places for the men and on six places for the women. The KNSB explicitly made clear that under all circumstances they want to decide on the final selection. Moreover, the KNSB claimed the freedom to fill out the tenth position on the PM ranking for a skater that is expected to increase the win probability on the team pursuit. By doing so, the calculated selection was used as a benchmark for the final KNSB selection, in the sense that the calculated win probabilities were leading in the discussion for the final selection.

In 2006, the year of the Torino Olympics, all first and second ranked OQT skaters with a qualification were selected. The remaining positions were filled in by using a combination of results of the previous Olympics and the last World Championship Single Distances. Using such 'old' results is obviously not very relevant for the determination of win probabilities of current skaters. The 2010 PM ranking methodology is based on recent performances, while the actual assignment of skaters to positions was communicated to the skaters and the coaches in an early stage and was completely transparent. Of course, not all skaters and coaches were completely satisfied, especially the 500m skaters felt themselves underestimated. Each selection system has it pros and cons, but the most important factor is clarity and support from the direct involved people. Athletes and coaches need to know when and how to qualify for the major tournaments so that they can tune their training schedules as early as possible and peak at the right moment.

### 4.5.1 Evaluations

It would be interesting to compare the PM rankings (see Table 4.12) with the actual results of the 2010 Olympic Games, such as to 'asses' the KNSB selection. This is done in Table 4.16. The second and fifth column contain the PM rankings of Table 4.15, the third and sixth column contain the KNSB selection, while the fourth and seventh column contain the results of the Olympic Games. During the Olympic

**Table 4.16.** Comparing PM ranking and Olympic results

| | **Men** | | | **Women** | | |
|---|---|---|---|---|---|---|
| | PM ranking | KNSB selection | Olympic ranking | PM ranking | KNSB selection | Olympic ranking |
| 1 | 10000m-nr1 | S. Kramer | 1 * | 1000m-nr1 | A. Gerritsen | 2 |
| 2 | 5000m-nr1 | S. Kramer | 1 | 500m-nr1 | A. Gerritsen | 4** |
| 3 | 1000mNr1 | S. Groothuis | 4 | 1000m-nr2 | L. van Riessen | 3 |
| 4 | 10000m-nr2 | B. de Jong | 4* | 1000m-nr3 | M. Boer | 6 |
| 5 | 5000m-nr2 | B. de Jong | 5 | 500m-nr2 | M. Boer | 5** |
| 6 | 1000m-nr2 | M. Tuitert | 5 | 1500m-nr1 | I. Wüst | 1 |
| 7 | 1500m-nr1 | M. Tuitert | 1 | 1500m-nr2 | M. Boer | 4 |
| 8 | 1500m-nr2 | S. Kuipers | 7 | 1000m-nr4 | I. Wüst | 8 |
| 9 | 1000m-nr3 | S. Kuipers | 6 | 500m-nr3 | T. Oenema | 16** |
| 10 | 10km-nr3 | A. v.d. Kieft | 10* | 1500m-nr3 | A. Gerritsen | 7 |
| 11 | 5km-nr3 | J. Blokhuizen | 9 | 3000m-nr1 | I. Wüst | 7 |
| 12 | 500m-nr1 | J. Smeekens | 6 | 5000m-nr1 | J. Voorhuis | 10 |
| 13 | 1500m-nr3 | S. Kramer | 13 | 3000m-nr2 | R. Groenewold | 10 |
| 14 | 1000m-nr4 | J. Bos | 12 | 5000m-nr2 | E. de Vries | 11 |
| 15 | 500m-nr2 | R. Mulder | 11 | 1500m-nr4 | L. van Riessen | 17 |
| 16 | 1500m-nr4 | S. Groothuis | 16 | 500m-nr4 | L. van Riessen | 20** |
| 17 | 500m-nr3 | S. Kuipers | 20 | 3000m-nr3 | D. Valkenburg | 11 |
| 18 | 500m-nr4 | J. Bos | 29 | | | |

The * and ** are the predicted and adjusted rankings due to the disqualifications of Kramer and Gerritsen.

Games two Dutch skaters were disqualified. Kramer was disqualified on the 10000m because of a wrong lane change. In order to make a fair assessment, we have given him the first position on the 10000m, mainly because he finished with the fastest time. Gerritsen was disqualified on her first 500m race, but we give her the fourth position based on her second 500m race. The other results are changed accordingly. The men's Olympic results match quite well with the KNSB PM ranking. Only the 1500m-nr1 and the 500m-nr1 are somewhat underestimated: the gold medal of Mark Tuitert on the 1500m came for many people as a surprise. However, from all Dutch skaters, Tuitert has the highest win-probability on the 1500m (see Table 4.3), which made him the most likely skater to win for the Netherlands. In case of the women, we see that all positions on 1500m are somewhat underestimated. The overestimation of the 500m-nr3 is caused by the fact that Marianne Timmer, with her excellent pre-Olympic season, could not participate.

In Table 4.17 we compared the KNSB PM ranking, the Calgary PM ranking and the Salt Lake City PM Ranking (see Table 4.12) with the Olympic results. The ordering of the eighteen (seventeen) distance labels in each PM ranking is compared with the ordering of the official outcome of the Olympic. The differences are measured by the Spearman's footrule (see Spearman (1904)), the normalized Spearman's footrule, and the Kendall rank correlation coefficient (see Kendall (1938)). Now, for two rankings $R1$ and $R2$ with the same $n$ objects, Spearman's footrule $D$ is defined as

$$D = \sum_{i=1}^{n} |r_{1i} - r_{2i}|$$

where $r_{1i}$ is the ranking position of object $i$ in $R1$. If $R1$ and $R2$ are identical, $D = 0$ and the maximum $D^{max}$ of $n^2/2$ if $n$ is even and $\frac{1}{2}(n+1)(n-1)$ if $n$ is odd. The normalized Spearman's footrule is defined by

$$ND = 1 - \frac{D}{D^{max}}.$$

Finally, the Kendall rank correlation coefficient is defined by

$$\tau = \frac{\sum_{i=1}^{n-1}\sum_{j=i}^{n} sgn(r_{1i} - r_{1j})sgn(r_{2i} - r_{2j})}{\frac{1}{2}n(n-1)}.$$

**Table 4.17.** Ranking correlation between PM ranking and Olympic outcome

| | **Men** | | | **Women** | | |
|---|---|---|---|---|---|---|
| | KNSB | Calgary | Salt Lake City | KNSB | Calgary | Salt Lake City |
| D | 20 | 22 | 26 | 34 | 36 | 40 |
| ND | 0.88 | 0.86 | 0.84 | 0.76 | 0.75 | 0.72 |
| $\tau$ | 0.84 | 0.83 | 0.80 | 0.72 | 0.72 | 0.66 |

D=Spearmans footrule, ND=Normalized Spearmans footrule, and $\tau$= Kendall rank correlation coefficient.

The results from Table 4.17 show a high correlation between the PM rankings and the actual Olympic results. This means that the PM rankings predict the ordering of distances on which the Netherlands will score quite well. It also shows that the final KNSB PM rankings scores the best. The small adjustments made by the KNSB (see Section 4.4.4) have resulted in a better PM ranking.

We also investigated whether or not Dutch skaters with the highest win probability also became the first Dutch Olympic skater. Table 4.18 lists for each distance the skater with the highest win probability, his rank on the Olympics, his rank among the Dutch skaters on the Olympics, and his rank during the OQT among the selected skaters. The table shows that all Dutch skaters with a highest win probability also

**Table 4.18.** Win probability comparison

| Men | | | | | |
|---|---|---|---|---|---|
| Distance | Skater with highest win prob | Medal win Prob. | Position Olympics | Among Dutch skaters | Ranking OQT |
| 500m | Jan Smeekens | 10% | 6 | 1 | 1 |
| 1000m | Stefan Groothuis | 77% | 4 | 1 | 1 |
| 1500m | Markt Tuitert | 17% | 1 | 1 | 3 |
| 5000m | Sven Kramer | 96% | 1 | 1 | 1 |
| 10000m | Sven Kramer | 100% | 1* | 1* | 1 |
| **Women** | | | | | |
| Distance | Skater with highest win prob | Medal win prob. | Position Olympics | Among Dutch skaters | Ranking OQT |
| 500m | Annette Gerritsen | 10% | 4** | 1 | 1 |
| 1000m | Annette Gerritsen | 69% | 2 | 1 | 1 |
| 1500m | Ireen Wüst | 29% | 1 | 1 | 1 |
| 3000m | Ireen Wüst | 1% | 7 | 1 | 1 |
| 5000m | Renate Groenewold | 0% | NS | - | - |

The * and ** are the predicted and adjusted rankings due to the disqualifications of Kramer and Gerritsen, NS = not started

became the highest classified Dutch Olympic skaters. The only exception is Renate Groenewold on the 5000m; although she has the highest win probability on this distance (mainly due to her result during the 2009 WSDCH), she was not selected for this distance. Mark Tuitert is the only skater with a highest win probability, but without winning the corresponding distance on the OQT.

# 4.6 Conclusions and future research

## 4.6.1 Fairness of the selection procedure

The selection of athletes for major tournaments, like the Olympics, is usually a very controversial matter. There are always disappointed athletes and coaches. In order to minimize the disappointments, the selection process needs to be fair, transparent and known long enough before the actual selection moment. The KNSB has decided, already in an early stage, to use a selection methodology with a high degree of objectiveness and transparency, and that was broadly accepted by skaters and coaches. The selection procedure needed to be known long enough before the December qualification tournament, because this tournament was the most decisive.

The Performance Matrices and Rankings, both based on recent performances of skaters, have served as major tools for deciding on the final selections. Because coaches and skaters were involved in the selection procedure from the start, only minor resistance appeared, certainly compared to the situation during the 2006 Torino selection procedure.

We have also calculated a theoretical 'optimal' selection, purely from the win probabilities. Although, the KNSB will always organize a selection tournament, as this creates a focus point for the skaters, the results show that the final KNSB selections do not differ very much from the computer calculated 'optimal' selections. This fact and the actual results from the Olympics indicate that the PM rankings are appropriate decision tools for selections.

We may therefore expect that in the future the KNSB will use a similar selection procedure, certainly as long as the ISU keeps the restriction of a maximum of ten men and ten women per country for all distances. The main complicating factor of the current selection procedure is the maximum of ten skaters per gender restriction over all events. This rule may result in a situation that the best athlete on a certain discipline needs to stay home in favor of an athlete with an higher win probability on another discipline. Therefore it is not a bad idea to drop this restriction and to use only restrictions per discipline.

### 4.6.2 Further improvements

For future use, we recommend a number of improvements. One of the drawbacks of the calculations is the low number of observations. Despite this fact, the ordered win probabilities in PM rankings do not differ much from the actual realizations during the Vancouver Winter Games, and besides a number of outliers, the selection order and the actual performance order match surprisingly close. But, of course, more data will make the results more robust.

Secondly, the way of calculating the individual performances of each skater could be changed. In stead of using the difference between the actual time and the average time of the five best skaters of each contest, an alternative would be to model the results of all tournaments as a mixed log-linear model (see Koning (2005)). The rink influences and seasonal influences are estimated as fixed effects and for the individual performance measure one may use the best linear unbiased estimator (BLUE) of the random effects. The advantage of such a model is that the performances in each tournament are no longer independent but correlated by the skaters. This gives the opportunity to deal with the fact that during some tournaments not all (best) skaters are present, whereas in our case the differences are biased by this fact. Furthermore, we may then use data from a longer period to obtain a robust estimation of the fixed effects and use the residuals of the recent tournaments to estimate the skaters effects and the corresponding win probabilities.

A second alternative is the interpretation of win probability. In this chapter we derive the win probabilities for individual skaters, but the selection order of distances is based only on the probability that a Dutch skater wins. Therefore in future research we also will look at win probabilities of countries and base the selection order on these probabilities. Questions like, ”what is the probability that a Dutch skater will win a gold medal?” should then be answered. But how to compare a skater with a win probability of 50% with four skaters that all have a win probability of 15% on another distance? Which distance should be put first in the PM ranking? Furhermore, the interpretation of the win probability of being within the best three becomes more complicated as it needs to incorperate the chance of winning more than one medal, as more skaters represent a country. Another improvement would be to include the win probabilities of the team pursuit.

Finally, the selection method can *mutatis mutandis* be used for other tournaments or other disciplines where there is a maximum number of participants for multiple events, such as, alpine skiing, and cross country skiing

Chapter 5

# Fairness and the 1000m Speed Skating

During the Olympic Games and the World Championships Single Distances the 1000m is skated by every skater only one time. However, there may be a difference in skating a 1000m race with a start in the inner and the outer lane that introduces an externality that introduces unfairness. We show that this difference indeed exists. For the period 2000-2009 we observe a statistically significant advantage of starting in the inner lane of 0.120 seconds for women. For male skater the difference between starting lanes is 0.030, but not significantly different from zero. In order to make the competition fair, we suggest that the 1000m should be skated twice. This Chapter is published as The Olympic 1000 Meter Speed Skating should be Skated Twice.

## 5.1 Introduction

The so-called 'uncertainty outcome' hypothesis in sports economics (see Fort (2006)) implies that fans value closeness of competition. Closeness of competition requires a balance of competition between teams, and between individual athletes. In sports, like speed skating, balance of competition ideally is the result of a more or less equal performance ability of the athletes. However, closeness of competition may also be a side effect of unfairness in the competition. This happens for instance if there is an advantage for one athlete over another one, because of the design of the competition. On the other hand, closeness of competition often results in smaller performance differences in timed events, like speed skating. If this difference is smaller than the error margin of the time registration system, the timed results may even be wrong. An example is the 1000m World Cup race between Simon Kuipers and Shani Davis in Calgary on November 18th 2007.

In this chapter we focus on the design of the 1000m speed skating competition, where there is a conflict between fairness of competition (a 'social' goal) and self-interest of individual speed skaters. In almost all economic models it is assumed that "... people are exclusively pursuing their material self-interest and do not care about 'social' goals per se" (see Frey and Schmidt (1999), p817). Although in economics this may be true for some people, in sports it is certainly true for all athletes: It is the essence of elite sports that all elite athletes are exclusively pursuing their material self-interest. Non cooperative behavior and inequality is the rule rather than the exception, and it sometimes involves cheating. Also the possibility of free riding (obviously within the rules of the game) is fully exploited. This article focuses on speed skating, and free riding in speed skating may occur if there is a lane advantage. This lane advantage—discussed in great detail below—constitutes an externality that is positive for one skater and negative for his or her opponent. To avoid free riding this behavior may be punished. However, since punishment can be costly to the punisher the self-interest hypothesis predicts zero punishment. In experimental economics this prediction is rejected. However, in sports it may hold true. A way to deal with market failures caused by externalities is through regulation. For instance, it may be decided to skate the 1000m twice with one start in the inner lane and one start in the outer lane. We analyze performance in sports events and particularly in speed skating, and investigate whether indeed there is a free riding in the 1000m speed skating.

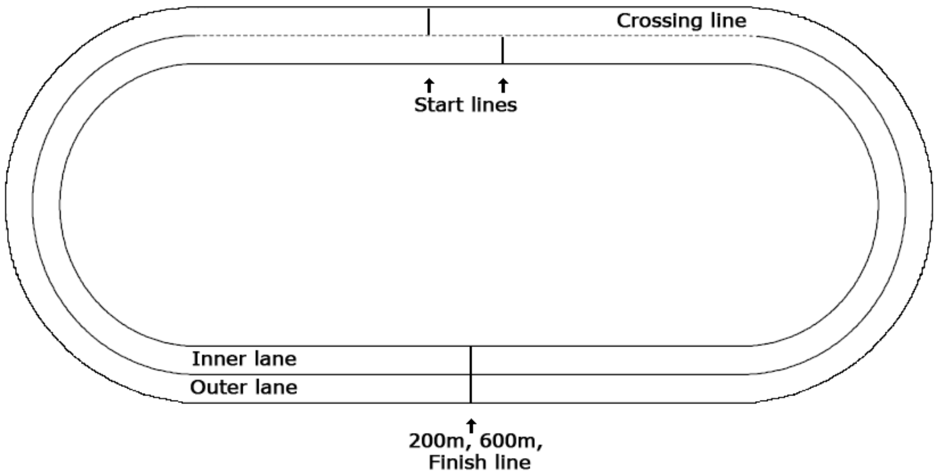
We focus on speed skating in which fairness of competition is facilitated by competing indoors, and by frequently cleaning and smoothing the ice with an ice resurfacer. However, in the 1000m speed skating event in which two skaters compete at the same time and switch lanes each lap, there may be a possible gain from starting in the inner lane in comparison with starting in the outer lane. Whether a skater starts in the inner lane or in the outer lane is determined by a pre-race draw. This unfairness may occur at the Olympic Games and the World Championships Single Distances in which the 1000m is skated by every skater only once.

In the next section we discuss the differences between starting in the inner lane and starting in the outer lane. In Section 5.3 we present a panel data model that accounts for unobserved individual specific effects (or unobserved heterogeneity). Examples of individual specific effects are a skater's length and weight. There may also be specific effects related to the rink, like the altitude of the speed skating oval, and whether a rink is covered or not. Technical improvements may be captured in seasonal effects. Section 5.4 discusses how we treat outliers that are due to slips and falls. In Section 5.6 we estimate the difference between starting in the inner lane and starting in the outer lane using the fixed effects model. We estimate separate effects for both male and female skaters. This section also tests the robustness of the results to removing outliers. Finally, Section 5.8 concludes.

## 5.2 Inner-outer lane differences

In a 1000m speed skating race on a 400m oval skaters race counter-clockwise in pairs, and start in a staggered position (see Figure 5.1). The staggered starting position is caused by the fact that the race of an inner lane starter has three inner lane curves and two outer lane curves, while a skater who starts in the outer lane completes three outer lane curves and two inner lane curves. When the skaters have completed the second curve, lanes are changed in a dedicated zone that is 70m in length. During the final 400m lap, there is a second change of lanes at the crossing line. Lap times are recorded after 200m, 600m, and 1000m. During a 1000m race there are several occasions that may result in a difference between 1000m times skated with a start in the inner and outer lane:

**Figure 5.1.** *Speed skating rink*




1. Due to the staggered starting positions, the outer lane starter starts in front of the inner lane starter, so that the inner lane starter can see and focus on his opponent;

2. Building up speed immediately after the start is easier in the shorter curve of the inner lane. However, it is more difficult to take the inner, more narrow, curves at top speed;

3. The inner lane starter has a disadvantage at the start, since he hears the starting shot later than the outer lane starter;

4. The change of lanes after passing the second curve may lead to a potential advantage for the outer lane starter, since he may make use of the slip stream of the inner lane starter during the change of lanes;

5. Similarly, a change of lanes at the crossing line after the fourth curve may lead to an advantage for the skater who starts in the inner lane, because he can make use of the slip stream of the other skater;

6. The skater who starts in the inner lane also finishes in the inner lane. This may lead to an advantage, since the longer curve of an outer lane is more difficult to skate when skaters are exhausted.

While some differences may be resolved, the difference between the number of inner and outer lane curves cannot be changed. So, there is still an unfair competition. We show that there is also a statistical difference in finishing times between 1000m races skated with a start in the inner and the outer lane. This inner-outer lane time difference is significant for female skaters. Our data set includes the results of the 1000m races for men and women of the Olympic Games (OG), World Single Distances Championships (WSDCh), World Sprint Championships (WSCh), and World Cup Competition (WCC) events between 2000 and 2009. Because of our findings, the Speed Skating Technical Committee of the International Skating Union had put our proposal on the agenda of the 53rd Ordinary Congress in Barcelona in 2010. Also our proposal to base the classification on the best time, and not on the total time of both races, has been on this agenda. Congress voted against both proposals. However, the discussion is still going on.

In the next section we present two versions of the statistical model: The fixed effects model and the random effects model. Both versions account for unobserved heterogeneity. Section 5.4 discusses how we treat outliers that are due to slips and falls. Section 5.5 tests whether the fixed effects or the random effects model is the preferred model. In Section 5.6 we estimate the difference between starting in the inner lane and starting in the outer lane using the preferred fixed effects model. We estimate separate effects for both male and female skaters. This section also tests the robustness of the results to removing outliers. Section 5.7 shows the implications of starting in the outer lane by calculating corrected 1000m rankings and comparing these rankings with the realized rankings. Finally, Section 5.8 concludes.

## 5.3 Statistical model

The analysis of performance in sports events is complicated. Many factors, such as training, nutrition, individual athletic abilities, like maximum oxygen intake in endurance events or muscular strength in sprint events, and technical progress in equipment determine the performance (see Atkinson and Nevill (2001), Kuper and Sterken (2003), and Kuper and Sterken (2008b)). Moreover, psychological and mental factors play a role. Together these factors determine the outcome of a fair competition in which the circumstances in which athletes compete are equal for all athletes.

In this section, we formulate a general model for estimating the difference in time between 1000m races skated with a start in the inner and the outer lane. The model is general in the sense that the abilities of individual skaters can be modeled by means of either a set of parameters or with a random variable, leading to either a fixed effects or a random effects model. We will explain both options, and describe a procedure to determine the option that models the abilities of the skaters correctly.

### 5.3.1 Definitions and data source

The definitions of the sets and the indices used to formulate the model are as follows:

$C$ = set of skaters in the dataset;

$J$ = set of nine seasons in the data set, $J = (2000\text{-}2001, \ldots, 2008\text{-}2009)$;

$K$ = set of 19 rinks in the data set (see Table 5.1);

$N_{cj}$ = set of contests organized during season $j\,(\in J)$ in which skater $c\,(\in C)$ participated;

$N_{ck}$ = set of contests skated at rink $k\,(\in K)$ in which skater $c\,(\in C)$ participated;

$N_c$ = set of contests in which skater $c\,(\in C)$ participated. Notice that $N_c = \cup_{j\in J} N_{cj}$, and $N_c = \cup_{k\in K} N_{ck}$ for each $c \in C$.

$n_c$ = number of contests in which skater $c\,(\in C)$ participated, i.e, $n_c = |N_c|$;

$N_c^I$ = set of contests in which skater $c\,(\in C)$ started in the inner lane;

$N_c^O$ = set of contests in which skater $c\,(\in C)$ started in the outer lane;

Finish$_{cijk}$ = 1000m finishing time of skater $c\,(\in C)$ skated on rink $k\,(\in K)$ during contest $i\,(\in N_c)$ in season $j\,(\in J)$.

**Table 5.1.** *Type and altitude (meters above sea level) of speed skating ovals*

| Location | Country | Type | Altitude |
|---|---|---|---|
| Changchun | China | Indoor | 210m |
| Berlin | Germany | Indoor | 34m |
| Calgary | Canada | Indoor | 1105m |
| Collalbo | Italy | Outdoor | 1173m |
| Erfurt | Germany | Indoor | 214m |
| Hamar | Norway | Indoor | 125m |
| Harbin | China | Indoor | 141m |
| Heerenveen | Netherlands | Indoor | 0m |
| Helsinki | Finland | Outdoor | 12m |
| Inzell | Germany | Outdoor | 691m |
| Kolomna | Russia | Indoor | 120m |
| Milwaukee | USA | Indoor | 216m |
| Moscow | Russia | Indoor | 127m |
| Nagano | Japan | Indoor | 346m |
| Oslo | Norway | Outdoor | 92m |
| Vancouver (Richmond) | Canada | Indoor | 4m |
| Salt Lake City | USA | Indoor | 1423m |
| Seoul | South Korea | Indoor | 63m |
| Torino | Italy | Indoor | 233m |

Source: *http://en.wikipedia.org/wiki/Speed_skating_rink*

With the exception of the second 2002-2003 WCC organized in Heerenveen, of which the results could not be traced, data of all men and women 1000m races of the OG, WCSD, WSC, and WC events in the period 2000-2009 are collected using the official website of the International Skating Union (www.isu.org). The first and second day of the WSC are considered as two different contests, just as the WCC events where the 1000m is skated twice. Data is collected from 108 different contests. The number of male and female skaters in our data set is 189 and 172, respectively. In total there are more than 5000 observations for men and women combined.

The data set is organized in such a way that the finishing times of one skater are included in one cluster. So the data set has a panel data structure, which can be used to estimate cluster-specific effects models (see Cameron and Trivedi (2007)). Because of injuries, selection procedures, and skaters ending their career in the period 2000-2009, none of the skaters participated in all contests of our data set. Moreover, the number of contests in which athletes start is different across skaters. Consequently, the panel considered is unbalanced.

### 5.3.2 General model

The events are organized on different ice rinks—some fast, some slow—and the events cover a period of nine skating seasons. So, we have to correct for the fact that performances are realized during different seasons and on different rinks. Let $c \in C$, $i \in N_c$, $j \in J$, and $k \in K$. We introduce two sets of dummies representing seasons and rinks:

$$\text{Season}_{cij} = \begin{cases} 1 & \text{if } i \in N_{cj} \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\text{Rink}_{cik} = \begin{cases} 1 & \text{if } i \in N_{ck} \\ 0 & \text{otherwise.} \end{cases}$$

The dummy that reflects information about the starting position (inner or outer lane) is denoted by $V_{ci}$, which is defined as:

$$V_{ci} = \begin{cases} 1 & \text{if } i \in N_c^O \\ 0 & \text{if } i \in N_c^I. \end{cases}$$

The dummies are included in the regression, together with a constant and an skater specific intercept, yielding the following general model:

$$\text{Finish}_{cijk} = \alpha + \theta_c + \sum_{j=1}^{9} \text{Season}_{cij}\beta_j + \sum_{k=1}^{19} \text{Rink}_{cik}\gamma_k + V_{ci}\delta + \epsilon_{cijk}, \tag{5.1}$$

with

$\alpha =$ constant term;

$\theta_c =$ skater specific constant that measures the ability of skater $c$: A low/high value of the ability parameter $\theta_c$ corresponds to the case in which skater $c$ performs better/worse than an average skater;

$\beta_j$ = parameter that measures the average speed of skaters during season $j$;

$\gamma_k$ = parameter that measures the average speed of skaters at rink $k$;

$\delta$ = parameter that measures the average advantage that an inner lane starter has over a skater starting in the outer lane during a 1000m race.

The term $\epsilon_{cijk}$ in Model (5.1) represents the error between the actual and estimated 1000m times of the $i$th contest of skater $c$ during season $j$ on rink $k$. It is assumed that for all skaters it holds that performances for all seasons $j$ and all rinks $k$ are correlated, and that the performances of different skaters are uncorrelated, i.e., for any $c_1, c_2 \in C$ with $c_1 \neq c_2$, $j_1, j_2 \in N_{c_1}$ and $j_3 \in N_{C_2}$ we assume that $\mathrm{Cov}(\epsilon_{j_1c_1}, \epsilon_{j_2c_1}) \neq 0$ and $\mathrm{Cov}(\epsilon_{j_1c_1}, \epsilon_{j_3c_2}) = 0$. The error terms are modeled in such a way that they are correlated within one cluster. In fact, we use the error terms to control for the dependence of performances over time.

The inclusion all season and rink dummies in the regression leads to an identification problem which is solved by removing some dummies. Since the choice of the dummies to remove is not relevant for the outcomes, we choose to remove the dummies corresponding to the season 2000-2001, and the rink in Berlin. The model then reads,

$$\mathrm{Finish}_{cijk} = \tilde{\alpha} + \theta_c + \sum_{l=2}^{9} \mathrm{Season}_{cil}\tilde{\beta}_l + \sum_{m=2}^{19} \mathrm{Rink}_{cim}\tilde{\gamma}_m + V_{ci}\delta + \epsilon_{cijk}, \tag{5.2}$$

with $\tilde{\alpha} = \alpha + \beta_1 + \gamma_1$, $\tilde{\beta}_l = \beta_l - \beta_1$, and $\tilde{\gamma}_m = \gamma_m - \gamma_1$. In model (5.2), the season parameter $\tilde{\beta}_l$ measures the average speed during season $l$ compared to the season 2000-2001, while the rink parameter $\tilde{\gamma}_m$ measures the average speed on rink $m$ compared to the rink in Berlin.

### 5.3.3 Random effects model and fixed effects model

Different assumptions on $\theta_c$ lead to quite different models. In a random effects model, the intercept $\theta_c$ is purely random, i.e., $\theta_c \sim \left(0, \sigma_\theta^2\right)$. An important assumption made in a random effects model is that the individual effect $\theta_c$ is uncorrelated with the regressors. When we would include a number of variables that explain the variation among the skaters, it is more likely that the individual effect is correlated with the regressors. Examples are a skater's length, weight, maximal oxygen intake and anaerobic threshold. According to Cameron and Trivedi (2007), violation of this assumption leads to biased estimates.

The difference between a random effects and a fixed effects model is that the latter does not require the assumption of uncorrelated regressors and individual effects. In a fixed effects model, the regression intercept $\theta_c$ is estimated by using a set of dummies, each dummy corresponding to a cluster. On the other hand, the individual effects can be removed by taking time averages in Model (5.2), and subtract the result obtained from Model (5.2), leading to the so-called within model (see Cameron and Trivedi (2007)).

### 5.3.4 Estimators

The random effects model and the within model are estimated with standard procedures in STATA. The vector of parameters

$$\phi = \left(\tilde{\beta}_2 \ \cdots \ \tilde{\beta}_9 \ \tilde{\gamma}_2 \ \cdots \tilde{\gamma}_{19} \ \delta\right)',$$

is estimated by using ordinary least squares. The standard errors of the estimates are corrected for heteroskedasticity and autocorrelation of the error terms. Estimates of the individual specific intercepts are denoted by $\widehat{\tilde{\alpha}} + \widehat{\theta}_c$:

$$\widehat{\tilde{\alpha}} + \widehat{\theta}_c = \overline{\text{Finish}}_c - \overline{X}_c \widehat{\phi}, \tag{5.3}$$

where $\overline{X}_c$ is the matrix containing the time averages of the time variant regressors from Model (5.2), namely

$$\overline{X}_c = \left(\overline{\text{Season}}_{c2} \ \cdots \ \overline{\text{Season}}_{c9} \ \overline{\text{Rink}}_{c2} \ \cdots \ \overline{\text{Rink}}_{c19} \ \overline{V}_c\right),$$

where

$$\overline{\text{Season}}_{cj} = \frac{1}{|N_{cj}|} \sum_{i \in N_c} \text{Season}_{icj}, \overline{\text{Rink}}_{ck} = \frac{1}{|N_{ck}|} \sum_{i \in N_c} \text{Rink}_{ick}, \ \overline{V}_c = \frac{1}{n_c} \sum_{i \in N_c} V_{ci},$$

for $c \in C$, $i \in N_c$, $j \in J \setminus \{1\}$, and $k \in K \setminus \{1\}$. The parameters with a hat ($\hat{}$) in Equation (5.3) refer to estimated values. Note that the all-unit vector $\iota$ is not a vector of the matrix $X_c$, since it is a time-invariant regressor.

### 5.3.5 Mundlak's model

The choice between a random effects and a fixed effects model hinges on the assumption of uncorrelated regressors and individual effects. We refer to Mundlak (1978) for a detailed discussion of the method that is used for testing whether or not this assumption holds. In fact, the set of regressors of Model (5.1) needs to be extended with the time averages of the variables in Model (5.1). As the variables are again perfect collinear, leading to an identification problem (see Section 5.3.2), we remove the variables $\text{Season}_{ic1}$, $\text{Rink}_{ic1}$, as well as the time averages of these variables, together with the all-unit vector of size $\sum_{c \in C} n_c$, so that Mundlak's model reads as follows,

$$\text{Finish}_{icjk} = \widetilde{XM}_{ic}\tilde{\rho} + \overline{\widetilde{XM}}_c \tilde{\eta} + \omega_c + \epsilon_{icjk}, \tag{5.4}$$

where

$$\widetilde{XM}_{ic} = \left(1 \ \text{Season}_{ic2} \cdots \ \text{Season}_{ic9} \ \text{Rink}_{ic2} \cdots \ \text{Rink}_{ic19} \ V_{ic}\right),$$

$$\overline{\widetilde{XM}}_c = \left(\overline{\text{Season}}_{c2} \ \cdots \ \overline{\text{Season}}_{c9} \ \overline{\text{Rink}}_{c2} \ \cdots \ \overline{\text{Rink}}_{c19} \ \overline{V}_c\right),$$

$$\tilde{\rho} = \left(\tilde{\alpha}' \ \tilde{\beta}_2 \ \cdots \ \tilde{\beta}_9 \ \tilde{\gamma}_2 \ \cdots \ \tilde{\gamma}_{19} \ \delta\right)', \text{ and } \tilde{\eta} = \left(\overline{\tilde{\beta}}_2 \ \cdots \ \overline{\tilde{\beta}}_9 \ \overline{\tilde{\gamma}}_2 \ \cdots \ \overline{\tilde{\gamma}}_{19} \ \overline{\delta}\right)'.$$

As in Section 5.3.2, the parameters of Model (5.4) satisfy: $\tilde{\beta}_j = \beta_j - \beta_1$, $\tilde{\gamma}_k = \gamma_k - \gamma_1$, $\bar{\tilde{\beta}}_j = \overline{\beta}_j - \overline{\beta}_1$, $\bar{\tilde{\gamma}}_k = \overline{\gamma}_k - \overline{\gamma}_1$, and $\tilde{\alpha}' = \alpha + \beta_1 + \gamma_1 + \overline{\alpha} + \overline{\beta}_1 + \overline{\gamma_1}$.

In Models (5.1) and (5.2) the individual effect is $\theta_c$, whereas in Model (5.4) this is $\widetilde{XM}_c\tilde{\eta} + \omega_c$ with $\omega_c$ being a random variable following the distribution $\omega_c \sim (0, \sigma_\omega^2)$. Note that $\tilde{\eta} = 0$ implies that the explanatory variables are uncorrelated with the individual effect $\theta_c$. The test of uncorrelated regressors and individual regression terms boils down to testing the null hypothesis $\mathrm{H}_0 : \tilde{\eta} = 0$ against the alternative $\mathrm{H}_A : \tilde{\eta} \neq 0$.

## 5.4 Detecting outliers

Speed skating is a technical sport that includes falls and minor slips, leading to results that deviate from normal performances. It is necessary to eliminate these outliers before the value of the parameter $\delta$ is estimated. In order to eliminate races with falls and slips, we introduce bounds on the 1000m finishing times, together with an outlier test. In Section 5.6.2 we show that our results are robust to deleting outliers.

### 5.4.1 Box plots

By introducing bounds on the 1000m finishing times, based on box plots of the $\text{Finish}_{icjk}$ values, we remove races containing a fall or slip. However, box plots of pooled finishing times neglect technological progress of equipment, and the fact that races are skated on different rinks. For example, during the season 2001-2002, Salt Lake City hosted the OG, while WC's were organized in Salt Lake City but also on outdoor low-altitude ovals, like the one in Oslo. Drawing conclusions based on box plots of the pooled data of the season 2001-2002 would lead to the removal of a considerable amount of finishing times (see Figure 5.2).

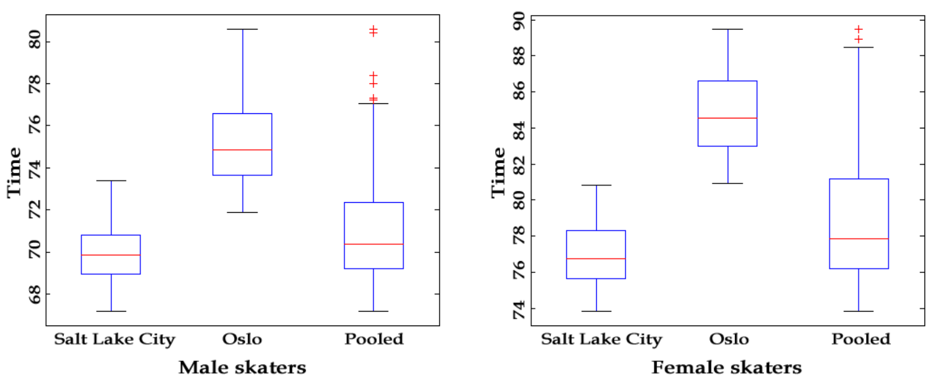



**Figure 5.2.** *Box plots of the finishing times of 2001-2002 events organized in Salt Lake City and Oslo. The third box plot in both panels corresponds to pooled finishing times. (+ is outlier)*

There are several outliers in the third box plot in both panels of Figure 5.2, although the first two box plots in both panels contain no outliers. Therefore, we construct box plots of 1000m times that are skated on the same rink and during the same season. If a finishing time is an outlier in a particular box plot, the corresponding bound is lowered and the finishing time is removed. This procedure is repeated until there are no outliers left.

### 5.4.2 Outlier test

In this section we use the same test statistic for the detection of outliers as in Hjort (2004) and Kamst *et al.* (2010) Let $c \in C$, $i \in N_c$, $j \in J$, and $k \in K$. The test statistic reads

$$T_{cijk} = \frac{\text{Finish}_{cijk} - \widehat{\text{Finish}}_{cijk}}{\sqrt{\widehat{\text{Var}}\left(\text{Finish}_{cijk}\right)}}, \tag{5.5}$$

where $\widehat{\text{Finish}}_{cijk}$ denotes the $i$th estimated 1000m time of skater $c$ on rink $k$ during season $j$, and is defined by,

$$\widehat{\text{Finish}}_{cijk} = \widehat{\widehat{\alpha}} + \widehat{\theta}_c + \sum_{j=2}^{9} \text{Season}_{cij}\widehat{\widehat{\beta}}_j + \sum_{k=2}^{19} \text{Rink}_{cik}\widehat{\widehat{\gamma}}_k + V_{ci}\widehat{\delta}. \tag{5.6}$$

Note that parameters with a hat (ˆ) denote estimated values.

In Equation (5.5), $\widehat{\text{Var}}\left(\text{Finish}_{cijk}\right)$ represents the estimated variance of the finishing time corresponding to the $i$th race of skater $c$ at location $k$ in season $j$. As mentioned above, we assume that the variance of the error term, $\sigma_\epsilon^2$, may differ across individuals. However, estimating the variance of this error term for each individual separately is not possible, because there are not enough observations for each skater. Therefore, we assume that the variance of the error term is constant across skaters, so that the usual estimator of the variance of the error terms can be used.

The 1000m time of a male skater is defined to be an outlier if the value of $T_{cijk}$ exceeds 2.60; for female skaters this value equals 2.80. A negative value of the test statistic corresponds to a good performance, suggesting that the 1000m time skated is not an outlier. An optimal choice of the bound values is a trade-off between a loss of information, caused by removing skaters from the starting lists, and the use of, what could be called, "bad" 1000 times. The latter could be a potential problem, since there may be large fluctuations among the estimates for different values of the bounds. We will address this problem in detail in Section 5.6.2.

## 5.5 Testing the random effects model

Both the random effects model and Mundlak's model are estimated. The latter model is used for testing whether or not the assumption of uncorrelated regressors and individual effects holds. To that end, we estimate the model based on the data set that is obtained after removing outliers based on box plots, and not on test statistics. Table 5.2 gives the estimates of both models.

**Table 5.2.** *Estimated parameters of the random effects and the Mundlak Model, using data of tournaments between 2000-2009. Par.=Parameter, Est.=Estimate, S.E.=Standard Error, and R.E.=Random Effects.*

| | | | **Men** | | | | **Women** | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | **R.E.** | | **Mundlak** | | **R.E.** | | **Mundlak** | |
| **Par.** | **Regressor** | | **Est.** | **S.E.** | **Est.** | **S.E.** | **Est.** | **S.E.** | **Est.** | **S.E.** |
| $\tilde{\alpha}$ | $\iota$ | Constant | 72.842 | 0.138 | - | - | 80.222 | 0.221 | - | - |
| $\tilde{\alpha}'$ | $\iota$ | Constant | - | - | 64.128 | 2.121 | - | - | 78.736 | 1.884 |
| $\tilde{\beta}_2$ | $\text{Season}_2$ | 2001-2002 | -0.023 | 0.083 | -0.038 | 0.083 | 0.104 | 0.144 | 0.089 | 0.145 |
| $\tilde{\beta}_3$ | $\text{Season}_3$ | 2002-2003 | -0.044 | 0.108 | -0.067 | 0.108 | 0.053 | 0.168 | 0.041 | 0.171 |
| $\tilde{\beta}_4$ | $\text{Season}_4$ | 2003-2004 | -0.216 | 0.133 | -0.243 | 0.134 | 0.211 | 0.224 | 0.194 | 0.228 |
| $\tilde{\beta}_5$ | $\text{Season}_5$ | 2004-2005 | -0.541 | 0.116 | -0.555 | 0.117 | -0.265 | 0.223 | -0.285 | 0.229 |
| $\tilde{\beta}_6$ | $\text{Season}_6$ | 2005-2006 | -0.217 | 0.117 | -0.327 | 0.115 | 0.341 | 0.233 | 0.208 | 0.239 |
| $\tilde{\beta}_7$ | $\text{Season}_7$ | 2006-2007 | -1.089 | 0.112 | -1.093 | 0.116 | -0.462 | 0.221 | -0.480 | 0.229 |
| $\tilde{\beta}_8$ | $\text{Season}_8$ | 2007-2008 | -1.127 | 0.120 | -1.127 | 0.125 | -0.557 | 0.237 | -0.584 | 0.246 |
| $\tilde{\beta}_9$ | $\text{Season}_9$ | 2008-2009 | -1.666 | 0.128 | -1.641 | 0.136 | -1.131 | 0.238 | -1.132 | 0.248 |
| $\tilde{\gamma}_2$ | $\text{Rink}_2$ | Calgary | -1.721 | 0.086 | -1.737 | 0.087 | -2.050 | 0.126 | -2.043 | 0.131 |
| $\tilde{\gamma}_3$ | $\text{Rink}_3$ | Changchun | 0.995 | 0.152 | 0.984 | 0.153 | 1.438 | 0.213 | 1.435 | 0.216 |
| $\tilde{\gamma}_4$ | $\text{Rink}_4$ | Collalbo | -0.821 | 0.110 | -0.746 | 0.110 | -0.082 | 0.17 | -0.050 | 0.173 |
| $\tilde{\gamma}_5$ | $\text{Rink}_5$ | Erfurt | -0.332 | 0.092 | -0.329 | 0.093 | -0.624 | 0.144 | -0.601 | 0.149 |
| $\tilde{\gamma}_6$ | $\text{Rink}_6$ | Hamar | -0.743 | 0.088 | -0.770 | 0.089 | -0.967 | 0.143 | -0.963 | 0.149 |
| $\tilde{\gamma}_7$ | $\text{Rink}_7$ | Harbin | 0.224 | 0.081 | 0.234 | 0.082 | 0.548 | 0.125 | 0.541 | 0.131 |
| $\tilde{\gamma}_8$ | $\text{Rink}_8$ | Heerenveen | -0.415 | 0.082 | -0.414 | 0.082 | -0.519 | 0.126 | -0.501 | 0.130 |
| $\tilde{\gamma}_9$ | $\text{Rink}_9$ | Helsinki | 2.906 | 0.190 | 2.905 | 0.191 | 4.705 | 0.233 | 4.706 | 0.236 |
| $\tilde{\gamma}_{10}$ | $\text{Rink}_{10}$ | Inzell | 1.001 | 0.107 | 0.977 | 0.107 | 1.629 | 0.152 | 1.600 | 0.155 |
| $\tilde{\gamma}_{11}$ | $\text{Rink}_{11}$ | Kolomna | -0.324 | 0.115 | -0.339 | 0.116 | -0.559 | 0.233 | -0.547 | 0.236 |
| $\tilde{\gamma}_{12}$ | $\text{Rink}_{12}$ | Milwaukee | -1.523 | 0.128 | -1.420 | 0.128 | -1.569 | 0.167 | -1.439 | 0.168 |
| $\tilde{\gamma}_{13}$ | $\text{Rink}_{13}$ | Moscow | 0.023 | 0.112 | -0.008 | 0.118 | -0.201 | 0.178 | -0.233 | 0.183 |
| $\tilde{\gamma}_{14}$ | $\text{Rink}_{14}$ | Nagano | -0.014 | 0.083 | -0.027 | 0.083 | 0.000 | 0.118 | 0.002 | 0.124 |
| $\tilde{\gamma}_{15}$ | $\text{Rink}_{15}$ | Oslo | 3.733 | 0.299 | 3.744 | 0.301 | 5.651 | 0.322 | 5.656 | 0.326 |
| $\tilde{\gamma}_{16}$ | $\text{Rink}_{16}$ | Salt Lake City | -2.211 | 0.082 | -2.214 | 0.082 | -2.568 | 0.131 | -2.558 | 0.136 |
| $\tilde{\gamma}_{17}$ | $\text{Rink}_{17}$ | Seoul | 1.286 | 0.101 | 1.233 | 0.099 | 2.135 | 0.150 | 2.083 | 0.154 |
| $\tilde{\gamma}_{18}$ | $\text{Rink}_{18}$ | Torino | -0.914 | 0.123 | -0.794 | 0.121 | -0.858 | 0.195 | -0.704 | 0.193 |
| $\tilde{\gamma}_{19}$ | $\text{Rink}_{19}$ | Vancouver | -0.027 | 0.136 | -0.030 | 0.136 | -0.220 | 0.181 | -0.203 | 0.184 |
| $\tilde{\delta}$ | $V$ | I/O | 0.025 | 0.025 | 0.026 | 0.025 | 0.118 | 0.033 | 0.123 | 0.033 |
| $\tilde{\bar{\beta}}_2$ | $\overline{\text{Season}_2}$ | $\overline{\text{2001-2002}}$ | - | - | 1.897 | 0.988 | - | - | 3.799 | 1.283 |
| $\tilde{\bar{\beta}}_3$ | $\overline{\text{Season}_3}$ | $\overline{\text{2002-2003}}$ | - | - | 1.983 | 0.961 | - | - | 2.162 | 1.131 |
| $\tilde{\bar{\beta}}_4$ | $\overline{\text{Season}_4}$ | $\overline{\text{2003-2004}}$ | - | - | 2.041 | 0.923 | - | - | 2.274 | 1.194 |
| $\tilde{\bar{\beta}}_5$ | $\overline{\text{Season}_5}$ | $\overline{\text{2004-2005}}$ | - | - | 0.652 | 1.208 | - | - | 3.384 | 1.334 |

*To be*

*continued*

*Table 5.2, continued*

| | | | Men | | | | Women | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | R.E. | | Mundlak | | R.E. | | Mundlak | |
| **Par.** | **Regressor** | | **Est.** | **S.E.** | **Est.** | **S.E.** | **Est.** | **S.E.** | **Est.** | **S.E.** |
| $\overline{\tilde{\beta}}_6$ | $\overline{\text{Season}_6}$ | $\overline{\text{2005-2006}}$ | - | - | 3.309 | 0.790 | - | - | 5.926 | 1.078 |
| $\overline{\tilde{\beta}}_7$ | $\overline{\text{Season}_7}$ | $\overline{\text{2006-2007}}$ | - | - | 1.845 | 0.819 | - | - | 2.369 | 1.102 |
| $\overline{\tilde{\beta}}_8$ | $\overline{\text{Season}_8}$ | $\overline{\text{2007-2008}}$ | - | - | 1.115 | 0.803 | - | - | 4.053 | 1.237 |
| $\overline{\tilde{\beta}}_9$ | $\overline{\text{Season}_9}$ | $\overline{\text{2008-2009}}$ | - | - | -0.168 | 0.999 | - | - | 1.492 | 2.002 |
| $\overline{\tilde{\gamma}}_2$ | $\overline{\text{Rink}_2}$ | $\overline{\text{Calgary}}$ | - | - | 8.297 | 2.148 | - | - | -2.331 | 2.173 |
| $\overline{\tilde{\gamma}}_3$ | $\overline{\text{Rink}_3}$ | $\overline{\text{Changchun}}$ | - | - | 10.167 | 2.631 | - | - | -3.273 | 2.884 |
| $\overline{\tilde{\gamma}}_4$ | $\overline{\text{Rink}_4}$ | $\overline{\text{Collalbo}}$ | - | - | -4.917 | 2.761 | - | - | -3.190 | 1.578 |
| $\overline{\tilde{\gamma}}_5$ | $\overline{\text{Rink}_5}$ | $\overline{\text{Erfurt}}$ | - | - | 2.881 | 2.530 | - | - | -10.482 | 2.886 |
| $\overline{\tilde{\gamma}}_6$ | $\overline{\text{Rink}_6}$ | $\overline{\text{Hamar}}$ | - | - | 8.001 | 2.099 | - | - | -0.777 | 1.498 |
| $\overline{\tilde{\gamma}}_7$ | $\overline{\text{Rink}_7}$ | $\overline{\text{Harbin}}$ | - | - | 1.996 | 2.360 | - | - | 0.203 | 1.610 |
| $\overline{\tilde{\gamma}}_8$ | $\overline{\text{Rink}_8}$ | $\overline{\text{Heerenveen}}$ | - | - | 8.230 | 2.255 | - | - | -1.685 | 1.477 |
| $\overline{\tilde{\gamma}}_9$ | $\overline{\text{Rink}_9}$ | $\overline{\text{Helsinki}}$ | - | - | 6.482 | 3.037 | - | - | 0.277 | 4.105 |
| $\overline{\tilde{\gamma}}_{10}$ | $\overline{\text{Rink}_{10}}$ | $\overline{\text{Inzell}}$ | - | - | 8.390 | 2.113 | - | - | 0.885 | 2.032 |
| $\overline{\tilde{\gamma}}_{11}$ | $\overline{\text{Rink}_{11}}$ | $\overline{\text{Kolomna}}$ | - | - | 12.135 | 2.746 | - | - | 2.693 | 3.527 |
| $\overline{\tilde{\gamma}}_{12}$ | $\overline{\text{Rink}_{12}}$ | $\overline{\text{Milwaukee}}$ | - | - | -0.116 | 3.613 | - | - | -21.846 | 4.441 |
| $\overline{\tilde{\gamma}}_{13}$ | $\overline{\text{Rink}_{13}}$ | $\overline{\text{Moscow}}$ | - | - | 9.802 | 2.183 | - | - | 1.749 | 0.914 |
| $\overline{\tilde{\gamma}}_{14}$ | $\overline{\text{Rink}_{14}}$ | $\overline{\text{Nagano}}$ | - | - | 8.353 | 2.127 | - | - | -1.434 | 1.685 |
| $\overline{\tilde{\gamma}}_{15}$ | $\overline{\text{Rink}_{15}}$ | $\overline{\text{Oslo}}$ | - | - | 2.623 | 3.215 | - | - | -1.606 | 1.903 |
| $\overline{\tilde{\gamma}}_{16}$ | $\overline{\text{Rink}_{16}}$ | $\overline{\text{Salt Lake City}}$ | - | - | 7.273 | 2.104 | - | - | -1.913 | 1.483 |
| $\overline{\tilde{\gamma}}_{17}$ | $\overline{\text{Rink}_{17}}$ | $\overline{\text{Seoul}}$ | - | - | 10.157 | 2.140 | - | - | 3.615 | 2.017 |
| $\overline{\tilde{\gamma}}_{18}$ | $\overline{\text{Rink}_{18}}$ | $\overline{\text{Torino}}$ | - | - | 2.190 | 2.268 | - | - | -7.087 | 2.303 |
| $\overline{\tilde{\gamma}}_{19}$ | $\overline{\text{Rink}_{19}}$ | $\overline{\text{Vancouver}}$ | - | - | 8.767 | 3.152 | - | - | -9.109 | 6.761 |
| $\overline{\delta}$ | $\overline{V}$ | $\overline{\text{I/O}}$ | - | - | -0.624 | 0.487 | - | - | -0.446 | 0.696 |

Actually, the Mundlak Model (5.4) is an extension of Model (5.2), obtained by adding the time averages of the regressors of Model (5.2) to the set of regressors. So the number of regressors in Model (5.4) is larger than in Model (5.2). The second column of Table 5.2 contains the regressors from Model (5.2) and Model (5.4), and the third column shows what the regressors refer to. Note that the regressor $\iota$, corresponding to the constants, is an all-unit vector. Computing the test statistic corresponding to test the null hypothesis $\text{H}_0 : \tilde{\eta} = 0$ against the alternative $\text{H}_A : \tilde{\eta} \neq 0$, the so-called Wald test (see Cameron and Trivedi (2007)), yields values of 320.60 and 199.37 for men and women, respectively. The Wald statistic has a Chi-square distribution with 27 degrees of freedom, because we are testing whether or not 27 parameters are significantly different from zero simultaneously. Obviously, the null hypothesis is rejected in both cases. This implies that the assumption of the random effects model that regressors are uncorrelated with the individual specific effect is not valid. So, we conclude that the individual effects should be modeled with a fixed effects model.

# 5.6 Estimation results

In this section we present the estimation results from the fixed effects model, and we test whether there is a difference between starting in the inner or in the outer lane. To that end, we estimate the parameter $\delta$ in Model (5.2). In order to test the influence of removing outliers on the results, we apply different values of the bound corresponding to the test statistics. We also give estimates of $\delta$ for different bound values.

## 5.6.1 Estimates of the fixed effects model

After removing times of races containing an error from the data set mentioned in Section 5.3.1, 2697 and 2529 observations for men and women, respectively, remain. The percentage of removed times of the original data set is 5.70 and 4.57, respectively. Table 5.3 presents the estimates of the parameters of Model (5.2), the robust standard errors, and the 95% confidence intervals.

**Table 5.3.** *Estimated parameters of the fixed effects model using data of tournaments between 2000-2009.*

| | | | Men | | | | Women | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Par.** | **Regressor** | | **Est.** | **S.E.** | **95% C.I.** | | **Est.** | **S.E.** | **95% C.I.** | |
| $\tilde{\beta}_2$ | $\text{Season}_2$ | 2001-2002 | -0.103 | 0.074 | -0.250 | 0.044 | 0.141 | 0.121 | -0.098 | 0.380 |
| $\tilde{\beta}_3$ | $\text{Season}_3$ | 2002-2003 | -0.165 | 0.092 | -0.347 | 0.016 | 0.138 | 0.145 | -0.148 | 0.425 |
| $\tilde{\beta}_4$ | $\text{Season}_4$ | 2003-2004 | -0.343 | 0.110 | -0.560 | -0.125 | 0.204 | 0.190 | -0.172 | 0.580 |
| $\tilde{\beta}_5$ | $\text{Season}_5$ | 2004-2005 | -0.607 | 0.102 | -0.808 | -0.406 | -0.213 | 0.196 | -0.601 | 0.175 |
| $\tilde{\beta}_6$ | $\text{Season}_6$ | 2005-2006 | -0.488 | 0.091 | -0.668 | -0.309 | 0.219 | 0.209 | -0.193 | 0.631 |
| $\tilde{\beta}_7$ | $\text{Season}_7$ | 2006-2007 | -1.104 | 0.100 | -1.302 | -0.906 | -0.433 | 0.200 | -0.828 | -0.038 |
| $\tilde{\beta}_8$ | $\text{Season}_8$ | 2007-2008 | -1.158 | 0.098 | -1.351 | -0.965 | -0.513 | 0.216 | -0.938 | -0.087 |
| $\tilde{\beta}_9$ | $\text{Season}_9$ | 2008-2009 | -1.701 | 0.121 | -1.940 | -1.461 | -1.116 | 0.213 | -1.536 | -0.696 |
| $\tilde{\gamma}_2$ | $\text{Rink}_2$ | Calgary | -1.775 | 0.082 | -1.937 | -1.614 | -2.032 | 0.109 | -2.247 | -1.818 |
| $\tilde{\gamma}_3$ | $\text{Rink}_3$ | Changchun | 0.989 | 0.160 | 0.674 | 1.304 | 1.529 | 0.192 | 1.150 | 1.908 |
| $\tilde{\gamma}_4$ | $\text{Rink}_4$ | Collalbo | -0.720 | 0.097 | -0.912 | -0.528 | 0.047 | 0.161 | -0.272 | 0.365 |
| $\tilde{\gamma}_5$ | $\text{Rink}_5$ | Erfurt | -0.358 | 0.089 | -0.534 | -0.181 | -0.576 | 0.130 | -0.833 | -0.319 |
| $\tilde{\gamma}_6$ | $\text{Rink}_6$ | Hamar | -0.825 | 0.085 | -0.992 | -0.658 | -0.902 | 0.133 | -1.165 | -0.640 |
| $\tilde{\gamma}_7$ | $\text{Rink}_7$ | Harbin | 0.212 | 0.080 | 0.055 | 0.370 | 0.595 | 0.117 | 0.364 | 0.825 |
| $\tilde{\gamma}_8$ | $\text{Rink}_8$ | Heerenveen | -0.455 | 0.078 | -0.609 | -0.301 | -0.440 | 0.111 | -0.660 | -0.220 |
| $\tilde{\gamma}_9$ | $\text{Rink}_9$ | Helsinki | 2.723 | 0.159 | 2.408 | 3.037 | 4.555 | 0.196 | 4.168 | 4.942 |
| $\tilde{\gamma}_{10}$ | $\text{Rink}_{10}$ | Inzell | 0.833 | 0.093 | 0.649 | 1.017 | 1.534 | 0.133 | 1.272 | 1.796 |
| $\tilde{\gamma}_{11}$ | $\text{Rink}_{11}$ | Kolomna | -0.353 | 0.115 | -0.581 | -0.126 | -0.452 | 0.216 | -0.878 | -0.025 |
| $\tilde{\gamma}_{12}$ | $\text{Rink}_{12}$ | Milwaukee | -1.333 | 0.115 | -1.559 | -1.106 | -1.346 | 0.157 | -1.655 | -1.037 |
| $\tilde{\gamma}_{13}$ | $\text{Rink}_{13}$ | Moscow | 0.026 | 0.112 | -0.194 | 0.247 | -0.128 | 0.152 | -0.427 | 0.172 |
| $\tilde{\gamma}_{14}$ | $\text{Rink}_{14}$ | Nagano | -0.109 | 0.074 | -0.254 | 0.037 | 0.036 | 0.103 | -0.168 | 0.239 |
| $\tilde{\gamma}_{15}$ | $\text{Rink}_{15}$ | Oslo | 2.604 | 0.176 | 2.257 | 2.951 | 5.071 | 0.316 | 4.447 | 5.695 |
| $\tilde{\gamma}_{16}$ | $\text{Rink}_{16}$ | Salt Lake City | -2.223 | 0.074 | -2.369 | -2.077 | -2.509 | 0.119 | -2.743 | -2.275 |
| $\tilde{\gamma}_{17}$ | $\text{Rink}_{17}$ | Seoul | 1.184 | 0.092 | 1.002 | 1.366 | 2.111 | 0.143 | 1.829 | 2.393 |
| $\tilde{\gamma}_{18}$ | $\text{Rink}_{18}$ | Torino | -0.702 | 0.105 | -0.910 | -0.494 | -0.607 | 0.187 | -0.976 | -0.238 |
| $\tilde{\gamma}_{19}$ | $\text{Rink}_{19}$ | Vancouver | -0.028 | 0.135 | -0.294 | 0.238 | -0.099 | 0.168 | -0.431 | 0.233 |
| $\delta$ | $V$ | I/O | 0.030 | 0.019 | -0.008 | 0.068 | 0.120 | 0.029 | 0.063 | 0.177 |

Par.=Parameter, Est.=Estimate, S.E.=Standard Error, and 95% C.I.=95% Confidence Interval.

Primarily, we are interested in the estimated value of $\delta$. Based on Table 5.3, we may conclude that there is a difference of 0.120 seconds between starting in the inner and the outer lane for female skaters, while this difference is 0.030 seconds for male skaters. These differences are significant, at a significance level of 95%, for women only.

There are more conclusions to be drawn from the estimation results. The gradual decline in the estimates of $\tilde{\alpha}_j$ $(j \in J \setminus \{1\})$, corresponding to the season dummies, confirm a positive development in finishing times. For male and female skaters, the estimates, corresponding to the season dummies of 2006-2007, 2007-2008, and 2008-2009, are significant, indicating that times skated during these seasons are significantly faster than in the season 2000-2001. For male skaters, this conclusion also holds for 2003-2004, 2004-2005, and 2005-2006. In contrast to indoor rinks, conditions regarding wind and temperature play a vital role on outdoor rinks. This fact is confirmed by the positive and significant estimated values of the parameters of the outdoor rink indicators, except for the high altitude rink in Collalbo. Moreover, the estimates of $\tilde{\beta}_2$ and $\tilde{\beta}_{16}$, corresponding to the ovals in Calgary and Salt Lake City, have negative values and are significant at 95%. This means that skaters are faster on these rinks than on the oval in Berlin because of the fact that Calgary and Salt Lake City are more than one kilometer above sea level (see Table 5.1).

### 5.6.2 Robustness with respect to outliers

Outliers may affect the estimated value of $\delta$, and other estimates. To test this we estimate Model (5.2) with different values of the bounds of the outlier test. The results are listed in Table 5.4.

**Table 5.4.** *Estimates of the difference between 1000m finishing times skated with a start in the inner and the outer lane for different values of the bound corresponding to the test statistic. S.E.$_\delta$=Standard Error of the estimate of $\delta$.*

| | **Men** | | **Women** | | | **Men** | | **Women** | |
|---|---|---|---|---|---|---|---|---|---|
| **Bound** | $\hat{\delta}$ | **S.E.**$_\delta$ | $\hat{\delta}$ | **S.E.**$_\delta$ | **Bound** | $\hat{\delta}$ | **S.E.**$_\delta$ | $\hat{\delta}$ | **S.E.**$_\delta$ |
| 4.60 | 0.026 | 0.024 | 0.123 | 0.033 | 3.50 | 0.026 | 0.021 | 0.115 | 0.033 |
| 4.50 | 0.031 | 0.023 | 0.123 | 0.033 | 3.40 | 0.030 | 0.021 | 0.118 | 0.033 |
| 4.40 | 0.031 | 0.023 | 0.126 | 0.032 | 3.30 | 0.027 | 0.021 | 0.118 | 0.033 |
| 4.30 | 0.029 | 0.022 | 0.126 | 0.032 | 3.20 | 0.023 | 0.020 | 0.117 | 0.033 |
| 4.20 | 0.029 | 0.022 | 0.126 | 0.032 | 3.10 | 0.029 | 0.020 | 0.121 | 0.031 |
| 4.10 | 0.032 | 0.022 | 0.123 | 0.032 | 3.00 | 0.020 | 0.020 | 0.126 | 0.030 |
| 4.00 | 0.029 | 0.022 | 0.127 | 0.032 | 2.90 | 0.021 | 0.021 | 0.115 | 0.029 |
| 3.90 | 0.025 | 0.022 | 0.124 | 0.032 | 2.80 | 0.022 | 0.020 | 0.120 | 0.029 |
| 3.80 | 0.025 | 0.022 | 0.118 | 0.033 | 2.70 | 0.023 | 0.020 | - | - |
| 3.70 | 0.025 | 0.022 | 0.118 | 0.033 | 2.60 | 0.030 | 0.019 | - | - |
| 3.60 | 0.030 | 0.021 | 0.118 | 0.033 | | | | | |

Without a formal test it is fair to conclude that—from comparing the standard errors with the estimates—the differences between estimates for different values of the bound are not significant. Hence, we conclude that the estimates of $\delta$ are robust. Eventually, we use test statistic bounds of 2.60 and 2.80 for male and female skaters, respectively. Decreasing the bound even more leads to the unnecessary removal of finishing times. In order to illustrate this, we analyze the results and the corresponding test values of the two skaters Rosendahl and Timmer (see Table 5.5).

**Table 5.5.** *Results of Rosendahl and Timmer in the 2001-2002 WSCh and 2008-2009 WC event in Nagano, respectively.*

**Men**

| Date | Location | Event | Name | Time | I/O |
|---|---|---|---|---|---|
| 20-1-2001 | Inzell | WSDCh | R. Rosendahl | 1:14.70 | O |
| 21-1-2001 | Inzell | WSDCh | R. Rosendahl | 1:14.75 | I |

**Women**

| Date | Location | Event | Name | Time | I/O |
|---|---|---|---|---|---|
| 13-12-2008 | Nagano | WCC | M. Timmer | 1:19.21 | I |
| 14-12-2008 | Nagano | WCC | M. Timmer | 1:19.23 | O |

Using the estimates that are obtained by using the data set without outliers for a bound value of 2.60, the test statistic in equation (5.5) of the second race of Rosendahl during the 2000-2001 WC event organized in Inzell is 2.56. While Rosendahl appears to be an outlier when a test statistic bound of 2.55 is used, Table 5.5 shows that there is a negligible difference between his 1000m times. For this reason, we argue that Rosendahl's second race does not contain a fall or slip, and that removing his second time from the data set is not necessary. Based on this argument, and the fact that the estimates are, in a statistical way, the same for bounds larger than 2.60, we use a bound of the test statistic of 2.60. A similar argument holds for the Timmer. The value of the test statistic is 2.72, but the difference between her first and second 1000m time is small (see Table 5.5). A statistical analysis of the characteristics of outliers is not meaningful, because the number of outliers is small, and would lead to imprecise inferences.

## 5.7 Correcting the 1000m times between 2000-2009

Table A.17 of the Appendix presents the actual and corrected rankings of the top five female skaters of all OG and WCSD events in the period 2000-2009, together with the ranking based on a fictitious reversed draw.

For all these events, the percentage of differences between the realized and corrected rankings is 29.06%. We observe considerable more differences between the actual ranking and the ranking based on a fictitious reversed draw, namely 48.29%. The only tournaments for which the top three finishers are identical in the three rankings of Table A.17 are the 2003 and 2008 WCSD, and the 2006 OG. More stricking is the fact that there are different gold medal winners in the rankings of the World Championships Single Distances of 2001, 2005, and 2009. For instance, in the WCSD of 2009, it turns out that Boer, starting in the outer lane, would have been second when we correct the actual ranking for her disadvantage of starting in the outer lane, while she would have won the gold medal if the draw of lanes would have been reversed.

## 5.8 Conclusion

Our investigations covering the period 2000-2009 show that female skaters starting and finishing in the inner lane have an advantage of $0.120$ seconds, on average, over their rivals finishing in the outer lane. This advantage is statistically significant. The 95% confidence interval for the difference in time between 1000m races skated with a start in the inner and the outer lane is $(0.063; 0.177)$. For men, the estimate of the parameter measuring the difference equals 0.030 seconds with a corresponding standard error of 0.019. The corresponding 95% confidence interval is given by $(-0.008; 0.068)$.

These estimates are robust with respect to deleting outliers, meaning that the estimates are not influenced by results that deviate from the average. This conclusion is based on sensitivity analysis with respect to the bound of the outlier tests. Tightening this bound leads to more outliers, but hardly affects the estimates. The results show that skaters starting in the inner lane have an advantage. This advantage, only for female skaters, is an example of free riding at the cost of one's opponent.

Besides the fact that inner and outer starting races are different races, the average significant difference in finishing times is large compared to the error margin of the time registration system. Especially for the top skaters the differences in finishing times are nowadays regularly within the error margins. Therefore we suggest that the 1000m is always skated twice during speed skating tournaments.

Based on a preliminary version of this paper, the Speed Skating Technical Committee of the International Skating Union has put our proposal on the agenda of the 53rd Ordinary Congress in Barcelona in 2010. Also our proposal to base the classification on the best time, and not on the total time of both races, is on this agenda. However, congress voted against both proposals. However, the discussion is still going on.

# Bibliography


Adler, N., Friedman, L., and Sinuany-Stern, Z. (2002), Review of ranking methods in the datat envelopment analysis context, *European Journal of Operational Research*, **140**, pp. 249–265.

Albright, S. C. (1993), A statistical analysis of hitting streask in baseball, *Journal of the American Statistical Assosiation*, **88**, pp. 1175–1183.

Atkinson, G. and Nevill, A. (2001), Selected issues in the design and analysis of sport performance research, *Journal of Sports Sciences*, **19**, pp. 811–827.

Berry, S. M., Reese, S., and Larkey, P. (1999), Bridging different eras in sports, *Journal of the American Statistical Assosiation*, **94**, pp. 661–687.

Bird, D. L. (2001), A brief history of ice and the national ice skating association of Great Britain, http://www.iceskating.org.uk/about/history.

Bosc, J. E., Liern, V., Martnez, A., and Sala, R. (2009), Increasing offensive or defensive efficiency? An analysis of Italian and Spanish football, *Omega*, **37**, pp. 63–78.

Brammer, C. L., Tanner, D. A., and Stager, J. M. (2009), Identification of bias in the natural porgression of swim performance, *Medicine and Science in Sport and Exercise*, **41**, p. S306.

Cameron, A. and Trivedi, P. (2007), *Microeconometrics*, Cambridge University Press.

Chatterjee, S. and Yilmaz, M. R. (1999), The NBA as an evolving multivariate system, *The American Statistican*, **53**, 3, pp. 257–262.

Cherchye, L. and Vermeulen, F. (2006), Robus rankings of mulitdimensional performances: an application to tour de france racing cyclists., *Journal of sports economics*, **7**, 4, pp. 359–373.

Dalby, A., Greve, J., and Jorsett, P. (2006), *Olympiske Vinterleker 1924 - 2006*, Akilles forlag.

Dumouchel, W. and O'Brien, F. (1991), Computing and graphics in statistics, chap. Integrating a robust option into a multiple regression computing environment, pp. 41–48, Springer-Verlag New York, Inc., New York, NY, USA.

Elandt, R. and Johnson, N. L. (1980), *Survival models and data analysis*, John Wiley and Sons.

Estellita-Lins, M. P., Gomes, E. G., Soares de Mello, J. C. C. B., and Soares de Mello, A. J. R. (2003), Olympic ranking based on a zero sum gains DEA model, *European Journal of Operational Research*, **148**, pp. 312–322.

Faugeras, J. (2011), European football clubs rankings, http://foot.faugeras.net/.

FIFA (2011), Fifa world ranking, http://www.fifa.com/ worldranking/rankingtable/index.html.

Fort, R. (2006), *Sports Economics*, Pearson Prentice Hall.

Frey, E. and Schmidt, K. M. (1999), A theory of fairness, competition, and cooperation, *Quarterly Journal of Economics*, **114**, pp. 817–868.

Galvão, R. D. and Raggi, L. A. (1989), A method for solving to optimality uncapacitated location problems, *Annals of Operations Research*, **18**, pp. 225–244.

Gould, S. (1996), *Full house: The spread of excellence from Plato to Darwin*, New York: Three Rivers Press.

Grubb, H. J. (1998), Models for comparing athletic performances, *The Statistician*, **47**, pp. 509–521.

Haas, D. J. (2003), Productive efficiency of english football teams. a data envelopment analysis approach., *Managerial and Decision Economics*, **24**, pp. 403–410.

Heiden, E. and Testa, M. (2008), *Faster, better, stronger*, Collins Living.

Heijmans (2001), Skate results, www.skateresults.com.

Hizen, Y. and Okui, R. (2009), Olympic athlete selection, *The Berkeley Electronic Journal of Economic Analysis and Policy*, **9**.

Hjort, N. L. (2004), Should the Olympic sprint skaters run the 500 meter twice, *T.B.P.*

Hlaoittinun, O., Bonjour, E., and Dulmet, M. (2007), A team buidling approach for competency development, *Industrial Engineering and Engineering Management*, **1**, pp. 1004 – 1008.

Houdijk, H., Koning, J. J. D., Groot, G. D., Bobbert, M. F., and Schenau, G. J. V. I. (2000), Push-off mechanics in speed skayting with conventionall skates and klapskates, *Medicine and Science in Sports and Exercise*, **32**, 3, pp. 635–641.

Hughes, M. and Franks, I. (2005), Analysis of passing sequensces, shots and goals in soccer, *Journal of Sport Sciences*, **23**, pp. 509–514.

Hughes, M. D. and Bartlett, R. M. (2002), The use of performance indicators in performance analysis, *Journal of Sport Sciences*, **20**, pp. 739–754.

Hughes, M. D., Hughes, M. T., and Behan, H. (2007), The evolution of computerised notational analysis through the example of racket sports, *International Journal of Sports Science and Engineering*, **1**, pp. 3–28.

Infostrada (2011), Euro club index: The ultimate club football ranking, www.euroclubindex.com.

International Association of Athletics Federation (2009), Drugs in sport/doping control, IAAF Medical Manual, www.iaaf.org.

ISU (2008), Qualification system for xxi winter Olympic Games Vancouver 2010, http://corporate.olympics.com.au/files/96/ISU_speedskating_sum_January08.pdf.

ISU (2010), Some key dates in ISU history, http://www2.isu.org/.

Kamst, R. (2010), Talent development tracking for speed skating: from high potentials to top performers, Master's thesis, University of Groningen.

Kamst, R., Kuper, G., and Sierksma, G. (2010), The Olympic 500 meter speed skating: The inner-outer lane difference., *Statistica Neerlandica*, **64**, pp. 448–459.

Kamst, R., Kuper, G. H., Sierksma, G., and Talsma, B. G. (2012), Inner-outer lane advantage in Olympic 1000m speed skating, *Journal of Economics and Statistics*, **232**, pp. 293–317.

Kendall, M. (1938), A new measure of rank correlation, *Biometrica*, **30**, pp. 81–89.

Kijkonderzoek (2011), http://www.kijkonderzoek.nl.

KNSB (2009), Selectienormen KNSB 2009-2010 langebaan voor deelname aan selectiewedstrijden, internationale wedstrijden, internationale kampioenschappen en Olympische winterspelen Vancouver.

KNSB (2010), KNSB strategisch plan, http://knsb.nl/publicaties/knsb-strategisch-plan/.

Koning, R. H. (2005), Home advantage in speed skating: evidence from individual data, *Journal of Sports Sciences*, **2**, 4, pp. 417–427.

Korver, E. and Hoogveld, A. (2008), Op zoek naar de juiste tijd, *SchaatSport*, **4**.

Kraemer, W. J. and Fleck, S. J. (2007), *Optimizing strength training: designing nonlinear periodization workouts*, Human Kinetics.

Kuper, G. and Sterken, E. (2003), Endurance in speed skating: the development of world records, *European Journal of Operational Research*, **148**, pp. 293–301.

Kuper, G. and Sterken, E. (2008a), *Statistical Thinking in Sports*, CrC Press.

Kuper, G. H. and Sterken, E. (2008b), Do skinsuits affect average skating speed?, *Sports Technology*, **1**, pp. 189–195.

Kuper, G. H. and Sterken, E. (2009), Do gender differences in ultimate elite cycling performance persist?

Lees, A. (2002), Technique analysis in sport: a critical review, *Journal of Sport Sciences*, **20**, pp. 813–828.

MATLAB (2010), *version 7.10.0 (R2010a)*, The MathWorks Inc, Natick, Massachusetts.

Milanovic, B. (2005), Globalization and goals: does soccer show the way?, *Review of International Political Economy.*, **12**, pp. 829–850.

Mundlak, Y. (1978), On the pooling of time series and cross section data, *Econometrica*, **46**, pp. 69–85.

Nevill, A., Atkinson, G., and Hughes, M. (2008), Twenty-five years of sport performance research in the Journal of Sport Sciences, *Journal of Sport Sciences*, **26**, pp. 413–426.

Nevill, A. M. and Holder, R. L. (1999), Home advantage in sport: an overview of studies on the advantage of playing at home, *Sports Medicine*, **28**, pp. 221–236.

Niemi, J. (2010), Evaluating individual player contribution in basketball, *JSM Proceedings, Section on Statistics in Sports. Vancouver, BC, Canada: American Statistical Association*, pp. 4914–4922.

NOS (2012), Tijdwaarneming bij veel sporten wankel, http://nos.nl/video/352905-tijdwaarneming-bij-veel-sporten-wankel.html.

Partridge, B. (2011), Fairness and performance enhancing swimsuits at the 2009 swimming world championships: the 'asterisk' championships, *Sport, Ethics and Philosophy*, **5**, pp. 63–74.

Reep, C., Pollard, R., and B.Benjamin (1971), Skill and chance in ball games, *Journal of the Ryal Statistical Society*, **134**, pp. 623–629.

Robinson, M. and Jonathan, A. (1995), Statistics for exeptional athletic records, *Applied Statistics*, **44**, 4, pp. 499–511.

Rodriguez, N. R., DiMarco, N. M., and Langley, S. (2009), Nutrition and athletic performance, *Medscape*.

RSSSF (2011), http://www.rsssf.com/intclub.html.

Ryvkin, D. (2010), The selection efficiency of tournaments, *European Journal of Operational Research*, **206**, pp. 667–675.

Schmidt, M. B. and Berri, D. J. (2003), On the evolution of competitive balance: The impact of an increasing global search, *Economic Inquiry*, **41**, 4, pp. 692–704.

Schmidt, M. B. and Berri, D. J. (2005), Concentration of playing talent. evolution in major league baseball, *Journal of Sport Economics*, **6**, pp. 412–419.

Schoonbeek, R. (2010), Dutch soccer players all-time ranking, Master's thesis, University of Groningen.

Scully, G. W. (2000), Diminshing returns and the limit of athletic performance, *Scotisch Journal of Political Economy*, **47**, pp. 456–470.

Sierksma, G. and Boon, B. (2003), Team formation: matching quality supply and quality demand, *European Journal of Operational Research*, **148**, pp. 272–292.

Snoep, H. (2004), *Statistiek Splitsen*, **10**, SchaatsSport.

Soleimani-Damaneh, J., Hamidi, M., and Sajadi, N. (2011), Evaluating the performance of iranian football teams utilizing linear programming, *American Journal of Operation Research*, **1**, pp. 65–72.

Spearman, C. (1904), The proof and measurement of association between two things, *American Journal of Psychology*, **15**, pp. 72–101.

Stefani, R. T. (1997), Survey of the major world sports rating systems, *Journal of Applied Statistics*, **24**, pp. 635–647.

Tatem, A. J., Guarra, C. A., Atkinson, P. M., and Hay, S. I. (2004), Athletics: momenous sprint at the 2156 Olympics?, *Nature*, **431**.

Thibault, V., Guillaume, M., Berthelot, G., Helou, N. E., Schaal, K., Quinquis, L., Nassif, H., Tafflet, M., Escolano, S., Hermine, O., and Toussaint, J.-F. (2010), Women and men in sport performance: The gender gap has not evolved since 1983, *Journal of Sports Science and Medicine*, **9**, p. 9.

Trewin, C. B., Hopkins, W., and Pyne, D. B. (2004), Relationship between world-ranking and Olympic performance of swimmers, *Journal of Sport Sciences*, **22**, pp. 229–345.

UCI (2011), Points scale uci worldtour ranking 2011, http://www.uciworldtour.com/Modules.

UEFA (2011), National team coefficient ranking: Technical explanation, http://www.uefa.com /MultimediaFiles/ Download /uefa/Key Topics/90/14/57/901457_DOWNLOAD.pdf.

USADA (2012), Report on proceedings under the world anti -doping code and the USADA protocol, http://d3epuodzu3wuis.cloudfront.net/ReasonedDecision.pdf.

Welsh, A. H. (1987), The trimmed mean in the linear model, *The Annals of Statistics*, **15**, pp. 20–36.

Whipp, B. and Ward, S. (1992), Will women soon outrun men?, *Nature*, **325**, 5.

Wieting, S. G. (2000), Twilight of the hero in Tour de France, *International Review of the Sociology of Sport*, **35**, pp. 348–363.

Williams, A. M. and Reilly, T. (2000), Talent identification and development in soccer, *Journal of Sport Sciences*, **18**, pp. 657–667.

Wittkowski (2004), An objective scoring method for graded count variables with application to olympics and baseball, *The Rockefeller University, General Clinical Research Center.*

Wu, J., Liang, L., and Chen, Y. (2009), DEA game cross-efficiency approach to Olympic rankings, *Omega*.

Yeadona, M. and Challisa, J. (1994), The future of performance related sports biomechanics research, *Journal of Sports Sciences*, **12**.

Zwols, Y. and Sierksma, G. (2009), Training optimization for the decathlon, *Operations Research*, **57**, pp. 812–822.

# Appendices

# Appendix A

## A.1 Universal Speed Skating Rankings

**Table A.1.** Ranking list Men 500m

| Position | Name | Country | Best per. | — Best three/four years | | |
|---|---|---|---|---|---|---|
| 1 | **Yevgeni Grishin** | URS | -0.197 | 1957 | 1961 | **1956** |
| 2 | **Uwe-Jens Mey** | GER | -0.193 | 1990 | 1991 | **1992** |
| 3 | **Hiroyasu Shimizu** | JPN | -0.158 | 2001 | 2000 | **1998** |
| 4 | **Eric Heiden** | USA | -0.157 | 1979 | 1978 | **1980** |
| 5 | Jeremy Wotherspoon | CAN | -0.143 | 2008 | 1999 | **2002** |
| 6 | Hasse Borjes | SWE | -0.097 | 1973 | 1970 | **1972** |
| 7 | **Yevgeni Grishin** | URS | -0.092 | **1956** | **1960** | **1964** |
| 8 | **Erhard Keller** | FRG | -0.090 | **1972** | 1968 | 1971 |
| 9 | Dan Jansen | USA | -0.082 | 1986 | **1988** | **1994** |
| 10 | **Sergey Fokichev** | URS | -0.069 | 1987 | **1984** | **1988** |
| 11 | Frode Ronning | NOR | -0.065 | 1981 | 1978 | **1980** |
| 12 | Leo Linkovesi | FIN | -0.057 | 1973 | 1974 | **1972** |
| 13 | **Ken Henry** | USA | -0.046 | **1952** | 1948 | 1955 |
| 14 | **Yevgeni Kulikov** | URS | -0.043 | 1975 | **1976** | 1980 |
| 15 | Igor Zhelezovski | URS | -0.041 | 1985 | 1989 | **1992** |
| 16 | Sergey Klevchenya | URS | -0.039 | 1996 | 1997 | **1994** |
| 17 | Manabu Horii | JPN | -0.028 | 1997 | **1994** | 1995 |
| 18 | Akira Kuroiwa | JPN | -0.028 | 1983 | 1987 | **1988** |
| 19 | Valeri Muratov | URS | -0.026 | **1976** | 1973 | 1978 |
| 20 | **Kang-Seok Lee** | KOR | -0.018 | 2007 | 2011 | **2010** |

**Table A.2.** Ranking list Women 500m

| Position | Name | Country | Best per. | Best three/four years | | |
|---|---|---|---|---|---|---|
| 1 | Jenny Wolf | GER | -0.378 | 2008 | 2009 | **2010** |
| 2 | **Bonnie Blair** | USA | -0.345 | 1995 | 1987 | **1994** |
| 3 | **Catriona LeMay** | CAN | -0.323 | **1998** | 2001 | 1999 |
| 4 | **Karin Enke** | GDR | -0.237 | **1980** | 1986 | 1982 |
| 5 | **Sheila Young** | USA | -0.224 | 1973 | **1976** | **1972** |
| 6 | **Christa Rothenburger** | GDR | -0.209 | **1984** | **1988** | 1989 |
| 7 | Qiaobo Ye | CHN | -0.171 | 1993 | **1992** | 1991 |
| 8 | Leah Poulos | USA | -0.155 | 1979 | **1980** | **1976** |
| 9 | **Lyudmila Titova** | URS | -0.146 | 1970 | **1968** | 1975 |
| 10 | Beixing Wang | CHN | -0.144 | 2007 | 2009 | **2010** |
| 11 | Manli Wang | CHN | -0.126 | 2004 | **2006** | 2003 |
| 12 | **Sang-Hwa Lee** | KOR | -0.096 | **2010** | 2011 | 2009 |
| 13 | Monique Garbrecht | GER | -0.089 | 2003 | 2000 | **2002** |
| 14 | Natalya Petrusyova | URS | -0.087 | 1982 | 1981 | **1980** |
| 15 | Angela Stahnke | GDR | -0.035 | 1990 | 1989 | **1988** |
| 16 | **Svetlana Zhurova** | RUS | -0.020 | **2006** | 2000 | 1999 |
| 17 | Ruihong Xue | CHN | 0.014 | 1997 | **1994** | 1996 |
| 18 | Susan Auch | CAN | 0.027 | 1998 | 1995 | **1994** |
| 19 | Natalya Glebova | URS | 0.032 | **1984** | 1982 | 1983 |
| 20 | **Anne Henning** | USA | 0.034 | 1971 | **1972** | 1970 |

**Table A.3.** Ranking list Men 1000m

| Position | Name | Country | Best per. | Best three years | | |
|---|---|---|---|---|---|---|
| 1 | **Eric Heiden** | USA | -0.485 | 1979 | 1978 | **1980** |
| 2 | Igor Zhelezovski | URS | -0.234 | 1985 | 1993 | **1992** |
| 3 | **Gaetan Boucher** | CAN | -0.229 | **1984** | 1985 | 1982 |
| 4 | **Ard Schenk** | NED | -0.229 | **1972** | 1971 | 1970 |
| 5 | **Shani Davis** | USA | -0.178 | 2008 | **2010** | 2009 |
| 6 | Erben Wennemars | NED | -0.173 | 2004 | 2003 | **2006** |
| 7 | Jan Bos | NED | -0.147 | 1999 | **1998** | 2000 |
| 8 | Sergey Khlebnikov | URS | -0.134 | 1983 | **1984** | 1982 |
| 9 | Yun-Man Kim | KOR | -0.104 | 1995 | **1992** | 1996 |
| 10 | **Peter Mueller** | USA | -0.045 | 1976 | 1977 | 1979 |
| 11 | **Dan Jansen** | USA | -0.042 | **1994** | 1986 | 1988 |
| 12 | Uwe-Jens Mey | GER | -0.022 | 1990 | **1988** | 1991 |
| 13 | Sergey Klevchenya | URS | -0.016 | 1996 | **1994** | 2001 |
| 14 | Stefan Groothuis | NED | 0.015 | 2011 | **2010** | 2009 |
| 15 | Kyu-Hyeok Lee | KOR | 0.026 | 2007 | **2010** | 2008 |
| 16 | **Gerard van Velde** | NED | 0.030 | 2002 | 2003 | 2001 |
| 17 | Valeri Muratov | URS | 0.031 | 1973 | **1976** | 1975 |
| 18 | Hasse Borjes | SWE | 0.033 | 1974 | 1973 | **1972** |
| 19 | Hans van Helden | FRA | 0.045 | **1976** | 1974 | 1973 |
| 20 | Nick Thometz | USA | 0.053 | 1987 | 1986 | **1984** |

**Table A.4.** Ranking list Women 1000m

| Position | Name | Country | Best per. | Best three years | | |
|---|---|---|---|---|---|---|
| 1 | **Karin Enke** | GDR | -0.423 | 1987 | 1986 | 1984 |
| 2 | **Natalya Petrusyova** | URS | -0.310 | 1980 | 1982 | 1983 |
| 3 | **Christine Nesbitt** | CAN | -0.285 | 2011 | 2009 | 2010 |
| 4 | Anni Friesinger | GER | -0.258 | 2004 | 2008 | 2006 |
| 5 | **Christa Rothenburger** | GDR | -0.232 | 1989 | 1988 | 1985 |
| 6 | **Bonnie Blair** | USA | -0.195 | 1994 | 1989 | 1990 |
| 7 | Sheila Young | USA | -0.166 | 1975 | 1974 | 1976 |
| 8 | **Chris Witty** | USA | -0.127 | 1998 | 1996 | 2002 |
| 9 | Ireen Wust | NED | -0.102 | 2007 | 2011 | 2006 |
| 10 | Cindy Klassen | CAN | -0.083 | 2007 | 2003 | 2006 |
| 11 | Monique Garbrecht | GER | -0.078 | 2001 | 1999 | 1992 |
| 12 | Andrea Mitscherlich | GDR | -0.075 | 1985 | 1984 | 1987 |
| 13 | Jennifer Rodriguez | USA | -0.064 | 2004 | 2003 | 2002 |
| 14 | Tatyana Averina | URS | -0.063 | 1974 | 1976 | 1979 |
| 15 | Sylvia Burka | CAN | -0.061 | 1977 | 1976 | 1979 |
| 16 | Lyudmila Titova | URS | -0.058 | 1971 | 1968 | 1975 |
| 17 | Sabine Volker | GER | -0.032 | 2002 | 2001 | 1999 |
| 18 | Qiaobo Ye | CHN | -0.026 | 1992 | 1991 | 1994 |
| 19 | Franziska Schenk | GER | -0.017 | 1997 | 1998 | 1994 |
| 20 | **Marianne Timmer** | NED | -0.011 | 1999 | 2006 | 2004 |

**Table A.5.** Ranking list Men Sprint

| Position | Name | Country | Best per. | Best four years | | | |
|---|---|---|---|---|---|---|---|
| 1 | Eric Heiden | USA | -0.254 | 1979 | **1980** | 1978 | 1977 |
| 2 | Igor Zhelezovski | URS | -0.128 | 1985 | 1989 | 1986 | **1992** |
| 3 | Uwe-Jens Mey | GER | -0.093 | 1990 | 1991 | 1989 | **1988** |
| 4 | Gaetan Boucher | CAN | -0.080 | **1984** | 1985 | 1982 | 1979 |
| 5 | Dan Jansen | USA | -0.042 | 1986 | **1994** | **1988** | 1989 |
| 6 | Jeremy Wotherspoon | CAN | -0.028 | 2008 | 2003 | 2000 | **1998** |
| 7 | Hiroyasu Shimizu | JPN | -0.006 | 1996 | 2000 | 1999 | **1998** |
| 8 | Sergey Klevchenya | URS | 0.016 | 1996 | **1994** | 1997 | 1995 |
| 9 | Kyu-Hyeok Lee | KOR | 0.023 | 2011 | 2008 | 2007 | **2010** |
| 10 | Frode Ronning | NOR | 0.030 | 1981 | 1982 | 1979 | **1980** |
| 11 | Valeri Muratov | URS | 0.033 | 1976 | 1973 | **1972** | 1970 |
| 12 | Erhard Keller | FRG | 0.047 | **1972** | 1968 | 1973 | 1971 |
| 13 | Peter Mueller | USA | 0.060 | **1976** | 1977 | 1979 | 1974 |
| 14 | Jan Bos | NED | 0.062 | 1999 | 2000 | **1998** | 2008 |
| 15 | Akira Kuroiwa | JPN | 0.084 | 1987 | 1983 | 1986 | **1988** |
| 16 | Sergey Khlebnikov | URS | 0.086 | 1982 | **1984** | 1981 | 1980 |
| 17 | Erben Wennemars | NED | 0.090 | 2003 | 2004 | **1998** | **2002** |
| 18 | Joey Cheek | USA | 0.111 | 2006 | 2003 | 2005 | **2002** |
| 19 | Gerard van Velde | NED | 0.123 | 2003 | **2002** | **1992** | 2005 |
| 20 | Shani Davis | USA | 0.124 | 2007 | 2009 | 2011 | **2010** |

**Table A.6.** Ranking list Women Sprint

| Position | Name | Country | Best per. | Best three/four years | | | |
|---|---|---|---|---|---|---|---|
| 1 | Karin Enke | GDR | -0.234 | 1986 | 1987 | **1984** | 1983 |
| 2 | Bonnie Blair | USA | -0.209 | **1994** | 1987 | 1989 | **1990** |
| 3 | Natalya Petrusyova | URS | -0.150 | 1982 | **1980** | 1983 | 1981 |
| 4 | Christa Rothenburger | GDR | -0.140 | 1989 | **1988** | 1986 | 1984 |
| 5 | Jenny Wolf | GER | -0.116 | 2008 | 2007 | 2009 | **2010** |
| 6 | Catriona LeMay | CAN | -0.102 | **1998** | **2002** | 2001 | 1999 |
| 7 | Sheila Young | USA | -0.091 | 1973 | **1976** | 1975 | 1981 |
| 8 | Monique Garbrecht | GER | -0.065 | 2001 | 2003 | 2000 | **2002** |
| 9 | Leah Poulos | USA | -0.023 | **1976** | 1979 | **1980** | 1977 |
| 10 | Qiaobo Ye | CHN | 0.014 | **1992** | 1993 | 1991 | **1994** |
| 11 | Angela Stahnke | GDR | 0.033 | 1990 | 1989 | 1985 | **1994** |
| 12 | Sabine Volker | GER | 0.039 | 2001 | **2002** | 1997 | 1999 |
| 13 | Anni Friesinger | GER | 0.041 | **2006** | 2007 | 2008 | 2004 |
| 14 | Franziska Schenk | GER | 0.062 | 1997 | **1994** | 1998 | 1995 |
| 15 | Beixing Wang | CHN | 0.101 | 2009 | 2007 | 2008 | **2010** |
| 16 | Sylvia Burka | CAN | 0.102 | 1977 | 1979 | **1976** | 1973 |
| 17 | Beth Heiden | USA | 0.131 | 1978 | 1979 | **1980** | 1977 |
| 18 | Atje Keulen-Deelstra | NED | 0.131 | 1973 | 1970 | **1972** | 1974 |
| 19 | Monika Pflug | FRG | 0.137 | **1972** | 1973 | 1981 | 1982 |
| 20 | Chris Witty | USA | 0.161 | 2000 | 1996 | 1999 | **1998** |

**Table A.7.** Ranking list Men 1500m

| Position | Name | Country | Best per. | Best three years | | |
|---|---|---|---|---|---|---|
| 1 | **Ard Schenk** | NED | -0.374 | 1973 | **1972** | 1971 |
| 2 | **Eric Heiden** | USA | -0.329 | 1979 | 1978 | **1980** |
| 3 | **Clas Thunberg** | FIN | -0.292 | 1925 | 1931 | **1924** |
| 4 | **Adne Sondral** | NOR | -0.252 | 2000 | 1999 | **1998** |
| 5 | Oscar Mathisen | NOR | -0.232 | 1912 | 1914 | 1913 |
| 6 | Ids Postma | NED | -0.206 | 1997 | **1998** | 1999 |
| 7 | **Johann Olav Koss** | NOR | -0.201 | **1994** | 1991 | 1993 |
| 8 | Shani Davis | USA | -0.194 | 2005 | 2004 | **2006** |
| 9 | **Jan Egil Storholt** | NOR | -0.175 | **1976** | 1973 | 1981 |
| 10 | **Mark Tuitert** | NED | -0.158 | 2004 | **2010** | 2005 |
| 11 | Rintje Ritsma | NED | -0.154 | 1995 | 1997 | **1998** |
| 12 | **Kees Verkerk** | NED | -0.140 | 1966 | 1969 | **1968** |
| 13 | Jaap Eden | NED | -0.131 | 1896 | 1895 | 1893 |
| 14 | Roar Gronvold | NOR | -0.123 | 1974 | **1972** | 1973 |
| 15 | Wim van der Voort | NED | -0.118 | 1951 | 1952 | 1953 |
| 16 | Boris Stenin | URS | -0.117 | 1962 | **1960** | 1959 |
| 17 | Rudolf Gundersen | NOR | -0.113 | 1906 | 1904 | 1902 |
| 18 | **Yevgeni Grishin** | URS | -0.112 | 1956 | **1960** | 1957 |
| 19 | Ivar Ballangrud | NOR | -0.107 | 1930 | 1926 | **1936** |
| 20 | Oleg Bozhev | URS | -0.104 | 1986 | **1984** | 1987 |

**Table A.8.** Ranking list Women 1500m

| Position | Name | Country | Best per. | Best three years | | |
|---|---|---|---|---|---|---|
| 1 | **Karin Enke** | GDR | -0.365 | 1986 | **1984** | 1987 |
| 2 | **Cindy Klassen** | CAN | -0.339 | **2006** | 2005 | 2004 |
| 3 | **Anni Friesinger** | GER | -0.331 | 2008 | 2003 | **2002** |
| 4 | **Ireen Wust** | NED | -0.304 | 2011 | 2007 | **2010** |
| 5 | Gunda Kleemann | GER | -0.292 | 1991 | 1995 | **1992** |
| 6 | **Lidia Skoblikova** | URS | -0.267 | **1964** | 1963 | 1967 |
| 7 | Andrea Mitscherlich | GDR | -0.189 | 1983 | 1987 | **1988** |
| 8 | **Emese Hunyady** | AUT | -0.150 | 1993 | **1994** | 1991 |
| 9 | **Yvonne van Gennip** | NED | -0.142 | 1988 | 1987 | **1992** |
| 10 | **Galina Stepanskaya** | URS | -0.134 | 1977 | **1976** | 1973 |
| 11 | Kristina Groves | CAN | -0.114 | 2009 | **2010** | 2008 |
| 12 | Stien Kaiser | NED | -0.095 | 1967 | 1969 | **1972** |
| 13 | Natalya Petrusyova | URS | -0.077 | 1981 | 1982 | **1984** |
| 14 | Inga Artamonova | URS | -0.072 | 1958 | **1962** | 1961 |
| 15 | Atje Keulen-Deelstra | NED | -0.055 | 1974 | 1973 | **1972** |
| 16 | Tatyana Averina | URS | -0.036 | 1978 | 1975 | **1976** |
| 17 | **Dianne Holum** | USA | -0.029 | **1972** | 1971 | 1967 |
| 18 | **Kaija Mustonen** | FIN | -0.023 | 1968 | **1964** | 1966 |
| 19 | **Annie Borckink** | NED | -0.018 | **1980** | 1975 | 1977 |
| 20 | **Jacqueline Borner** | GER | -0.008 | **1992** | 1990 | 1989 |

**Table A.9.** Ranking list Men 5000m

| Position | Name | Country | Best per. | Best three years | | |
|---|---|---|---|---|---|---|
| 1 | **Gianni Romme** | NED | -0.621 | 2000 | **1998** | 2003 |
| 2 | **Hjalmar Andersen** | NOR | -0.553 | 1951 | **1952** | 1950 |
| 3 | **Johann Olav Koss** | NOR | -0.483 | **1994** | 1991 | 1993 |
| 4 | Jaap Eden | NED | -0.481 | 1895 | 1896 | 1893 |
| 5 | **Ard Schenk** | NED | -0.466 | **1972** | 1973 | 1970 |
| 6 | **Sven Kramer** | NED | -0.464 | 2008 | 2009 | **2010** |
| 7 | Peder Ostlund | NOR | -0.327 | 1899 | 1898 | 1900 |
| 8 | **Fred Anton Maier** | NOR | -0.326 | **1968** | 1967 | 1965 |
| 9 | **Viktor Kosichkin** | URS | -0.324 | **1960** | 1962 | 1961 |
| 10 | **Eric Heiden** | USA | -0.292 | 1979 | 1978 | **1980** |
| 11 | **Knut Johannesen** | NOR | -0.290 | 1963 | **1964** | 1959 |
| 12 | **Ivar Ballangrud** | NOR | -0.286 | 1926 | **1936** | 1932 |
| 13 | **Jochem Uytdehaage** | NED | -0.265 | 2002 | 2003 | **2006** |
| 14 | **Chad Hedrick** | USA | -0.260 | 2005 | **2006** | 2004 |
| 15 | **Sten Stensen** | NOR | -0.248 | 1975 | **1976** | 1973 |
| 16 | Kees Verkerk | NED | -0.229 | 1967 | 1970 | **1968** |
| 17 | **Geir Karlstad** | NOR | -0.205 | **1992** | 1987 | 1985 |
| 18 | **Clas Thunberg** | FIN | -0.189 | **1932** | **1924** | 1925 |
| 19 | Leo Visser | NED | -0.149 | 1989 | 1987 | **1992** |
| 20 | Oscar Mathisen | NOR | -0.144 | 1912 | 1914 | 1908 |

**Table A.10.** Ranking list Women 3000m

| Position | Name | Country | Best per. | Best three years | | |
|---|---|---|---|---|---|---|
| 1 | **Gunda Kleemann** | GER | -0.513 | 1991 | **1996** | **1992** |
| 2 | **Andrea Mitscherlich** | GDR | -0.429 | 1987 | **1984** | 1983 |
| 3 | **Martina Sablikova** | CZE | -0.403 | 2011 | **2010** | 2009 |
| 4 | Karin Enke | GDR | -0.363 | 1986 | **1984** | 1987 |
| 5 | **Stien Kaiser** | NED | -0.334 | **1972** | 1967 | 1969 |
| 6 | **Ans Schut** | NED | -0.310 | 1970 | **1968** | 1969 |
| 7 | **Bjorg Eva Jensen** | NOR | -0.243 | **1980** | 1979 | 1978 |
| 8 | **Claudia Pechstein** | GER | -0.239 | 2000 | 2003 | **1998** |
| 9 | **Lidia Skoblikova** | URS | -0.218 | 1960 | **1964** | 1963 |
| 10 | **Yvonne van Gennip** | NED | -0.203 | **1988** | 1987 | 1985 |
| 11 | Heike Schalling | GDR | -0.184 | 1993 | 1989 | **1992** |
| 12 | **Ireen Wust** | NED | -0.174 | 2011 | 2007 | **2006** |
| 13 | Anni Friesinger | GER | -0.166 | 2003 | 2005 | **2002** |
| 14 | Inga Artamonova | URS | -0.149 | 1958 | **1962** | 1961 |
| 15 | Beth Heiden | USA | -0.139 | 1979 | **1980** | 1977 |
| 16 | Renate Groenewold | NED | -0.114 | 2004 | 2007 | **2006** |
| 17 | Valentina Stenina | URS | -0.064 | 1966 | 1963 | **1960** |
| 18 | Cindy Klassen | CAN | -0.061 | **2006** | 2005 | 2007 |
| 19 | Karin Kessow | GDR | -0.011 | 1975 | **1976** | 1974 |
| 20 | Emese Hunyady | AUT | -0.005 | **1994** | **1992** | 1993 |

**Table A.11.** Ranking list Men 10000m

| Position | Name | Country | Best per. | Best three years | | |
|---|---|---|---|---|---|---|
| 1 | **Johann Olav Koss** | NOR | -0.686 | **1994** | 1991 | 1993 |
| 2 | **Gianni Romme** | NED | -0.649 | 2000 | **1998** | 1997 |
| 3 | **Hjalmar Andersen** | NOR | -0.618 | **1952** | 1951 | 1950 |
| 4 | **Jonny Nilsson** | SWE | -0.503 | 1963 | **1964** | 1962 |
| 5 | **Knut Johannesen** | NOR | -0.488 | 1957 | **1960** | 1959 |
| 6 | Viktor Kosichkin | URS | -0.477 | 1961 | **1960** | 1962 |
| 7 | **Bob de Jong** | NED | -0.466 | 2011 | 2003 | **2006** |
| 8 | **Tomas Gustafson** | SWE | -0.445 | **1988** | 1982 | 1983 |
| 9 | Sten Stensen | NOR | -0.424 | **1974** | 1976 | **1978** |
| 10 | **Ard Schenk** | NED | -0.391 | 1971 | 1973 | **1972** |
| 11 | **Jochem Uytdehaage** | NED | -0.314 | **2002** | 2004 | 2005 |
| 12 | **Bart Veldkamp** | BEL | -0.304 | 1993 | 1990 | **1992** |
| 13 | Geir Karlstad | NOR | -0.260 | 1985 | 1987 | **1992** |
| 14 | **Piet Kleine** | NED | -0.249 | **1976** | 1973 | 1977 |
| 15 | **Eric Heiden** | USA | -0.234 | 1979 | **1980** | 1977 |
| 16 | Sven Kramer | NED | -0.224 | 2007 | 2008 | **2006** |
| 17 | Leo Visser | NED | -0.210 | 1989 | **1988** | 1990 |
| 18 | Fred Anton Maier | NOR | -0.199 | **1968** | 1967 | 1969 |
| 19 | Kees Verkerk | NED | -0.197 | 1974 | 1966 | **1972** |
| 20 | **Ivar Ballangrud** | NOR | -0.174 | 1930 | 1929 | **1936** |

**Table A.12.** Ranking list Women 5000m

| Position | Name | Country | Best per. | Best three/four years | | |
|---|---|---|---|---|---|---|
| 1 | **Gunda Kleemann** | GER | -0.746 | 2001 | 2000 | **1998** |
| 2 | **Martina Sablikova** | CZE | -0.712 | 2007 | **2010** | 2008 |
| 3 | **Claudia Pechstein** | GER | -0.482 | **1998** | **2002** | 2000 |
| 4 | Andrea Mitscherlich | GDR | -0.473 | 1986 | 1987 | **1988** |
| 5 | **Yvonne van Gennip** | NED | -0.408 | **1988** | 1987 | 1992 |
| 6 | Stephanie Beckert | GER | -0.328 | 2011 | **2010** | 2009 |
| 7 | Gretha Smit | NED | -0.314 | 2004 | **2002** | 2003 |
| 8 | **Clara Hughes** | CAN | -0.220 | 2003 | 2005 | **2006** |
| 9 | Heike Schalling | GDR | -0.083 | 1993 | 1991 | **1992** |
| 10 | Gabi Schonbrunn | GDR | -0.076 | **1984** | 1988 | 1985 |
| 11 | Eevi Huttunen | FIN | -0.050 | 1951 | 1953 | 1954 |
| 12 | Cindy Klassen | CAN | -0.002 | **2006** | 2003 | 2001 |
| 13 | Rimma Zhukova | URS | 0.030 | 1955 | 1949 | 1952 |
| 14 | Carla Zijlstra | NED | 0.053 | 1997 | 1995 | **1992** |
| 15 | Karin Enke | GDR | 0.060 | 1987 | **1988** | 1984 |
| 16 | Elena Belci | ITA | 0.138 | **1996** | 1990 | **1994** |
| 17 | Sabine Brehm | GDR | 0.139 | 1986 | 1983 | **1984** |
| 18 | Anni Friesinger | GER | 0.145 | 2003 | 2005 | **2002** |
| 19 | Ireen Wust | NED | 0.195 | 2011 | 2008 | 2007 |
| 20 | Kristina Groves | CAN | 0.231 | 2008 | 2009 | **2006** |

**Table A.13.** Ranking list Men Overall

| Position | Name | Country | Best per. | Best four years | | | |
|---|---|---|---|---|---|---|---|
| 1 | Eric Heiden | USA | -0.263 | 1979 | 1978 | **1980** | 1976 |
| 2 | Ard Schenk | NED | -0.224 | **1972** | 1973 | 1971 | 1967 |
| 3 | Sven Kramer | NED | -0.214 | 2008 | 2009 | 2007 | **2010** |
| 4 | Johann Olav Koss | NOR | -0.183 | **1994** | 1991 | 1993 | 1990 |
| 5 | Oscar Mathisen | NOR | -0.153 | 1912 | 1914 | 1913 | 1908 |
| 6 | Jaap Eden | NED | -0.128 | 1896 | 1895 | 1893 | 1894 |
| 7 | Gianni Romme | NED | -0.126 | **1998** | 2000 | 2003 | 2002 |
| 8 | Hjalmar Andersen | NOR | -0.122 | 1951 | **1952** | 1950 | 1954 |
| 9 | Clas Thunberg | FIN | -0.098 | 1925 | **1924** | 1931 | 1929 |
| 10 | Ivar Ballangrud | NOR | -0.097 | 1930 | **1936** | 1926 | 1938 |
| 11 | Rintje Ritsma | NED | -0.044 | 1995 | **1998** | 1996 | 1993 |
| 12 | Ids Postma | NED | -0.006 | **1998** | 1996 | 1997 | 2001 |
| 13 | Kees Verkerk | NED | 0.007 | 1967 | 1969 | 1966 | **1968** |
| 14 | Oleg Goncharenko | URS | 0.013 | 1953 | 1958 | 1954 | **1956** |
| 15 | Bernt Evensen | NOR | 0.026 | 1927 | 1934 | **1928** | 1931 |
| 16 | Tomas Gustafson | SWE | 0.030 | **1988** | 1983 | 1982 | 1986 |
| 17 | Birger Wasenius | FIN | 0.053 | 1939 | 1934 | **1936** | 1937 |
| 18 | Michael Staksrud | NOR | 0.058 | 1935 | 1930 | 1937 | **1928** |
| 19 | Knut Johannesen | NOR | 0.065 | 1963 | **1964** | 1957 | 1955 |
| 20 | Roald Larsen | NOR | 0.065 | **1924** | 1926 | 1922 | 1925 |

**Table A.14.** Ranking list Women Overall

| Position | Name | Country | Best per. | Best four years | | | |
|---|---|---|---|---|---|---|---|
| 1 | Gunda Kleemann | GER | -0.411 | 1995 | 1991 | 1997 | **1998** |
| 2 | Karin Enke | GDR | -0.306 | 1980 | 1986 | 1987 | **1984** |
| 3 | Lidia Skoblikova | URS | -0.176 | 1963 | **1964** | **1960** | 1962 |
| 4 | Andrea Mitscherlich | GDR | -0.168 | 1983 | **1984** | 1985 | 1987 |
| 5 | Claudia Pechstein | GER | -0.161 | 2000 | 1998 | **1994** | 2001 |
| 6 | Anni Friesinger | GER | -0.112 | 2005 | **2008** | 2007 | 2002 |
| 7 | Natalya Petrusyova | URS | -0.091 | 1982 | 1981 | 1983 | **1980** |
| 8 | Inga Artamonova | URS | -0.058 | **1958** | 1965 | **1962** | 1957 |
| 9 | Cindy Klassen | CAN | -0.057 | **2006** | 2003 | 2005 | 2007 |
| 10 | Martina Sablikova | CZE | -0.028 | **2010** | 2009 | 2011 | 2008 |
| 11 | Bonnie Blair | USA | -0.022 | **1994** | **1992** | 1988 | 1986 |
| 12 | Ireen Wust | NED | -0.014 | 2007 | 2011 | 2008 | **2006** |
| 13 | Valentina Stenina | URS | -0.002 | 1961 | 1965 | 1966 | **1960** |
| 14 | Stien Kaiser | NED | 0.015 | 1967 | **1972** | 1965 | 1966 |
| 15 | Rimma Zhukova | URS | 0.018 | 1955 | **1956** | 1953 | 1954 |
| 16 | Yvonne van Gennip | NED | 0.026 | **1988** | 1985 | 1989 | 1992 |
| 17 | Sheila Young | USA | 0.029 | **1976** | 1975 | 1973 | 1974 |
| 18 | Tatyana Averina | URS | 0.056 | 1978 | **1976** | 1974 | 1979 |
| 19 | Atje Keulen-Deelstra | NED | 0.058 | 1973 | 1974 | **1972** | 1970 |
| 20 | Tamara Rylova | URS | 0.085 | 1959 | 1958 | 1956 | **1960** |

**Table A.15.** Ranking 4 vs 3 seasons

| Skater | Pos. 4s | Pos. 3s | Dif. | Score 4s | Score 3s |
|---|---|---|---|---|---|
| Gunda Kleemann | 1 | 1 | 0 | -0.411 | -0.414 |
| Karin Enke | 2 | 2 | 0 | -0.306 | -0.324 |
| Lidia Skoblikova | 3 | 3 | 0 | -0.176 | -0.209 |
| Andrea Mitscherlich | 4 | 5 | -1 | -0.168 | -0.184 |
| Claudia Pechstein | 5 | 4 | 1 | -0.161 | -0.185 |
| Anni Friesinger | 6 | 7 | -1 | -0.112 | -0.113 |
| Natalya Petrusyova | 7 | 8 | -1 | -0.091 | -0.097 |
| Inga Artamonova | 8 | 10 | -2 | -0.058 | -0.070 |
| Cindy Klassen | 9 | 6 | 3 | -0.057 | -0.114 |
| Martina Sablikova | 10 | 9 | 1 | -0.028 | -0.081 |
| Bonnie Blair | 11 | 11 | 0 | -0.022 | -0.054 |
| Ireen Wust | 12 | 13 | -1 | -0.014 | -0.029 |
| Valentina Stenina | 13 | 17 | -4 | -0.002 | -0.008 |
| Stien Kaiser | 14 | 14 | 0 | 0.015 | -0.017 |
| Rimma Zhukova | 15 | 16 | -1 | 0.018 | -0.009 |
| Yvonne van Gennip | 16 | 12 | 4 | 0.026 | -0.029 |
| Sheila Young | 17 | 15 | 2 | 0.029 | -0.012 |
| Tatyana Averina | 18 | 19 | -1 | 0.056 | 0.037 |
| Atje Keulen-Deelstra | 19 | 18 | 1 | 0.058 | 0.031 |
| Tamara Rylova | 20 | 22 | -2 | 0.085 | 0.092 |
| Beth Heiden | 21 | 20 | 1 | 0.113 | 0.077 |
| Sofya Kondakova | 22 | 23 | -1 | 0.138 | 0.103 |
| Lidia Selikhova | 23 | 26 | -3 | 0.149 | 0.120 |
| Kristina Groves | 24 | 25 | -1 | 0.161 | 0.117 |
| Christine Nesbitt | 25 | 31 | -6 | 0.181 | 0.172 |
| Nina Statkevich | 26 | 34 | -8 | 0.187 | 0.176 |
| Jennifer Rodriguez | 27 | 36 | -9 | 0.193 | 0.185 |
| Ans Schut | 28 | 30 | -2 | 0.193 | 0.172 |
| Irina Yegorova | 29 | 35 | -6 | 0.193 | 0.179 |
| Sylvia Burka | 30 | 29 | 1 | 0.199 | 0.166 |
| Leah Poulos | 31 | 24 | 7 | 0.200 | 0.114 |
| Dianne Holum | 32 | 28 | 4 | 0.219 | 0.160 |
| Randi Thorvaldsen | 33 | 33 | 0 | 0.225 | 0.174 |
| Renate Groenewold | 34 | 32 | 2 | 0.228 | 0.172 |
| Gabi Schonbrunn | 35 | 38 | -3 | 0.235 | 0.207 |
| Chris Witty | 36 | 39 | -3 | 0.239 | 0.215 |
| Anzhelika Kotyuga | 37 | 45 | -8 | 0.257 | 0.248 |
| Lyudmila Titova | 38 | 41 | -3 | 0.257 | 0.223 |
| Heike Schalling | 39 | 42 | -3 | 0.258 | 0.225 |
| Qiaobo Ye | 40 | 40 | 0 | 0.258 | 0.217 |
| Galina Stepanskaya | 41 | 37 | 4 | 0.261 | 0.196 |
| Marianne Timmer | 42 | 50 | -8 | 0.264 | 0.267 |
| Kaija Mustonen | 43 | 44 | -1 | 0.275 | 0.241 |
| Emese Hunyady | 44 | 43 | 1 | 0.281 | 0.235 |
| Jacqueline Borner | 45 | 49 | -4 | 0.300 | 0.265 |
| Carry Geijssen | 46 | 46 | 0 | 0.304 | 0.253 |
| Daniela Anschutz | 47 | 48 | -1 | 0.313 | 0.261 |
| Eevi Huttunen | 48 | 62 | -14 | 0.316 | 0.335 |
| Seiko Hashimoto | 49 | 59 | -10 | 0.326 | 0.321 |
| Annamarie Thomas | 50 | 57 | -7 | 0.330 | 0.319 |

Pos.= Position, s4= four seasons, s3= three seasons, Dif=Difference between Pos.s4 and Pos.s3

## A.2 Soccer Rankings

**Table A.16.** Top 10 of Dutch soccer players per decade

| **1956-1960** | | **1961-1970** | | **1971-1980** | |
|---|---|---|---|---|---|
| Tinus Bosselaar | 75 | Theo van Duivenbode | 405 | Ruud Krol | 1503 |
| Piet van der Kuil | 59 | Coen Moulijn | 340 | Arie Haan | 1471 |
| Piet Ouderland | 45 | Sjaak Swart | 302 | Johan Neeskens | 1402 |
| Eddy Pieters Graafland | 44 | Rinus Israël | 285 | Johan Cruijff | 1216 |
| Sjaak Swart | 33 | Johan Cruijff | 269 | Wim Suurbier | 1106 |
| Coen Dillen | 23 | Henk Groot | 264 | Johnny Rep | 950 |
| Noud van Melis | 15 | Piet Keizer | 256 | Rob Rensenbrink | 879 |
| Jan Klaassens | 15 | Henk Wery | 249 | Wim Jansen | 854 |
| Jan Notermans | 15 | Bennie Muller | 234 | Piet Keizer | 766 |
| Cor van der Hart | 14 | Eddy Pieters Graafland | 233 | Gerrie Mühren | 743 |
| **1981-1990** | | **1991-2000** | | **2001-2010** | |
| Marco van Basten | 931 | Frank de Boer | 933 | Mark van Bommel | 884 |
| Frank Rijkaard | 879 | Edwin van der Sar | 729 | Clarence Seedorf | 822 |
| Ruud Gullit | 674 | Ronald de Boer | 684 | Gio. van Bronckhorst | 717 |
| Gerald Vanenburg | 654 | Edgar Davids | 681 | Edwin van der Sar | 669 |
| Hans van Breukelen | 651 | Dennis Bergkamp | 646 | Dirk Kuijt | 666 |
| Berry van Aerle | 617 | Clarence Seedorf | 582 | Wesley Sneijder | 647 |
| Ronald Koeman | 600 | Michael Reiziger | 577 | Arjen Robben | 604 |
| Arnold Mühren | 518 | Ronald Koeman | 545 | Phillip Cocu | 475 |
| Jan Wouters | 470 | Danny Blind | 543 | Robin van Persie | 467 |
| Adri van Tiggelen | 442 | Marc Overmars | 540 | Ruud van Nistelrooij | 410 |
| **2011-2011** | | | | | |
| Edwin van der Sar | 144 | | | | |
| Wesley Sneijder | 41 | | | | |
| Rafael van der Vaart | 28 | | | | |
| Klaas-Jan Huntelaar | 28 | | | | |
| Theo Janssen | 22 | | | | |
| Robin van Persie | 21 | | | | |
| Gregory van der Wiel | 21 | | | | |
| Clarence Seedorf | 20 | | | | |
| Arjen Robben | 18 | | | | |
| Maarten Stekelenburg | 17 | | | | |

# A.3 Speed Skating 1000m results with correction

**Table A.17.** *Realized and corrected rankings of the OG and the WCSD in the period 2000-2009.*

**WCSD, Salt Lake City 2001**

| Realized ranking | | | | Corrected ranking | | | | Reversed draw ranking | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | M. Garbrecht | I | 1:14.13 | 1 | S. Völker | I | 1:14.02 | 1 | S. Völker | I | 1:14.02 |
| 2 | S. Völker | O | 1:14.14 | 2 | M. Garbrecht | I | 1:14.13 | 2 | M. Garbrecht | O | 1:14.25 |
| 3 | C. Lemay-Doan | I | 1:14.50 | 3 | C. Witty | I | 1:14.47 | 3 | C. Witty | I | 1:14.47 |
| 4 | C. Witty | O | 1:14.59 | 4 | C. Lemay-Doan | I | 1:14.50 | 4 | C. Lemay-Doan | O | 1:14.62 |
| 5 | A. Friesinger | O | 1:14.75 | 5 | A. Friesinger | I | 1:14.63 | 5 | A. Friesinger | I | 1:14.63 |

**OG, Salt Lake City 2002**

| Realized ranking | | | | Corrected ranking | | | | Reversed draw ranking | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | C. Witty | I | 1:13.83 | 1 | C. Witty | I | 1:13.83 | 1 | C. Witty | O | 1:13.95 |
| 2 | S. Völker | I | 1:13.96 | 2 | S. Völker | I | 1:13.96 | 2 | S. Völker | O | 1:14.08 |
| 3 | J. Rodriguez | I | 1:14.24 | 3 | J. Rodriguez | I | 1:14.24 | 3 | A. Friesinger | I | 1:14.35 |
| 4 | M. Timmer | I | 1:14.45 | 4 | A. Friesinger | I | 1:14.35 | 4 | J. Rodriguez | O | 1:14.36 |
| 5 | A. Friesinger | O | 1:14.47 | 5 | M. Timmer | I | 1:14.45 | 5 | M. Garbrecht | I | 1:14.48 |

**WCSD, Berlin 2003**

| Realized ranking | | | | Corrected ranking | | | | Reversed draw ranking | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A. Friesinger | O | 1:16.85 | 1 | A. Friesinger | I | 1:16.73 | 1 | A. Friesinger | I | 1:16.73 |
| 2 | J. Rodriguez | O | 1:17.28 | 2 | J. Rodriguez | I | 1:17.16 | 2 | J. Rodriguez | I | 1:17.16 |
| 3 | C. Klassen | O | 1:17.36 | 3 | C. Klassen | I | 1:17.24 | 3 | C. Klassen | I | 1:17.24 |
| 4 | M. Timmer | I | 1:17.96 | 4 | A. Kotyuga | I | 1:17.86 | 4 | A. Kotyuga | I | 1:17.86 |
| 5 | A. Kotyuga | O | 1:17.98 | 5 | M. Timmer | I | 1:17.96 | 5 | M. Garbrecht | I | 1:18.07 |

**WCSD, Seoul 2004**

| Realized ranking | | | | Corrected ranking | | | | Reversed draw ranking | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A. Friesinger | I | 1:17.82 | 1 | A. Friesinger | I | 1:17.82 | 1 | A. Friesinger | O | 1:17.94 |
| 2 | M. Timmer | I | 1:18.50 | 2 | M. Timmer | I | 1:18.5 | 2 | C. Klassen | I | 1:18.56 |
| 3 | C. Klassen | O | 1:18.68 | 3 | C. Klassen | I | 1:18.56 | 3 | M. Timmer | O | 1:18.62 |
| 4 | A. Kotyuga | O | 1:18.80 | 4 | A. Kotyuga | I | 1:18.68 | 4 | A. Kotyuga | I | 1:18.68 |
| 4 | J. Rodriguez | O | 1:18.80 | 4 | J. Rodriguez | I | 1:18.68 | 4 | J. Rodriguez | I | 1:18.68 |

**WCSD, Inzell 2005**

| Realized ranking | | | | Corrected ranking | | | | Reversed draw ranking | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | B. de Loor | I | 1:18.24 | 1 | B. de Loor | I | 1:18.24 | 1 | A. Friesinger | I | 1:18.34 |
| 2 | A. Friesinger | O | 1:18.46 | 2 | A. Friesinger | I | 1:18.34 | 2 | B. de Loor | O | 1:18.36 |
| 3 | M. Timmer | I | 1:18.71 | 3 | J. Rodriguez | I | 1:18.66 | 3 | J. Rodriguez | I | 1:18.66 |
| 4 | C. Klassen | I | 1:18.75 | 4 | M. Timmer | I | 1:18.71 | 4 | M. Timmer | O | 1:18.83 |
| 5 | J. Rodriguez | O | 1:18.78 | 5 | C. Klassen | I | 1:18.75 | 5 | C. Klassen | O | 1:18.87 |

**OG, Torino 2006**

| Realized ranking | | | | Corrected ranking | | | | Reversed draw ranking | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | M. Timmer | O | 1:16.05 | 1 | M. Timmer | I | 1:15.93 | 1 | M. Timmer | I | 1:15.93 |
| 2 | C. Klassen | I | 1:16.09 | 2 | C. Klassen | I | 1:16.09 | 2 | C. Klassen | O | 1:16.21 |
| 3 | A. Friesinger | I | 1:16.11 | 3 | A. Friesinger | I | 1:16.11 | 3 | A. Friesinger | O | 1:16.23 |
| 4 | I. Wüst | I | 1:16.39 | 4 | I. Wüst | I | 1:16.39 | 4 | I. Wüst | O | 1:16.51 |
| 5 | K. Groves | I | 1:16.54 | 5 | K. Groves | I | 1:16.54 | 5 | B. de Loor | I | 1:16.61 |

**WCSD, Salt Lake City 2007**

| Realized ranking | | | | Corrected ranking | | | | Reversed draw ranking | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | I. Wüst | I | 1:13.83 | 1 | I. Wüst | I | 1:13.83 | 1 | I. Wüst | O | 1:13.95 |
| 2 | A. Friesinger | O | 1:14.26 | 2 | A. Friesinger | I | 1:14.14 | 2 | A. Friesinger | I | 1:14.14 |
| 3 | C. Nesbitt | I | 1:14.44 | 3 | C. Simionato | I | 1:14.41 | 3 | C. Simionato | I | 1:14.41 |
| 4 | C. Simionato | O | 1:14.53 | 4 | C. Nesbitt | I | 1:14.44 | 4 | C. Nesbitt | O | 1:14.56 |
| 5 | F. Wang | O | 1:14.96 | 5 | F. Wang | I | 1:14.84 | 5 | F. Wang | I | 1:14.84 |

**WCSD, Nagano 2008**

| Realized ranking | | | | Corrected ranking | | | | Reversed draw ranking | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A. Friesinger | I | 1:15.37 | 1 | A. Friesinger | I | 1:15.37 | 1 | A. Friesinger | O | 1:15.49 |
| 2 | K. Groves | I | 1:16.01 | 2 | K. Groves | I | 1:16.01 | 2 | K. Groves | O | 1:16.13 |
| 3 | A. Gerritsen | O | 1:16.35 | 3 | A. Gerritsen | I | 1:16.23 | 3 | A. Gerritsen | I | 1:16.23 |
| 4 | C. Nesbitt | I | 1:16.44 | 4 | C. Nesbitt | I | 1:16.44 | 4 | C. Nesbitt | O | 1:16.56 |
| 5 | P. van Deutekom | I | 1:16.89 | 5 | P. van Deutekom | I | 1:16.89 | 5 | P. van Deutekom | O | 1:17.01 |

**WCSD, Vancouver 2009**

| Realized ranking | | | | Corrected ranking | | | | Reversed draw ranking | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | C. Nesbitt | I | 1:16.28 | 1 | C. Nesbitt | I | 1:16.28 | 1 | M. Boer | I | 1:16.32 |
| 2 | A. Friesinger | I | 1:16.32 | 2 | A. Friesinger | I | 1:16.32 | 2 | C. Nesbitt | O | 1:16.40 |
| 3 | M. Boer | O | 1:16.44 | 2 | M. Boer | I | 1:16.32 | 3 | A. Friesinger | O | 1:16.44 |
| 4 | N. Bruintjes | I | 1:16.80 | 4 | N. Bruintjes | I | 1:16.80 | 4 | Y. Lobysheva | I | 1:16.82 |
| 5 | S. Yoshii | I | 1:16.82 | 5 | S. Yoshii | I | 1:16.82 | 5 | L. van Riessen | I | 1:16.90 |
| 6 | P. Jin | I | 1:16.87 | 5 | Y. Lobysheva | I | 1:16.82 | 6 | N. Bruintjes | O | 1:16.92 |

# Samenvatting

Het centrale thema van dit proefschrift betreft het ontwikkelen van technieken om prestaties van topsporters te vergelijken. Een van de belangrijkste motieven om dat te doen is dat op basis van prestaties en de ontwikkeling ervan selectiebeslissingen moeten worden genomen. Daarbij wordt ook een aanzet gedaan om factoren te bepalen, die van invloed zijn op de verbetering van prestaties. Verder is ook gekeken naar de externe factoren, die tijdens de wedstrijden van invloed zijn op de prestaties.

Er is wellicht geen bedrijfstak waar getallen en statistieken een grotere rol spelen dan de sport. Prestaties worden geleverd, gemeten, ge-analyseerd en vergeleken. Daarnaast is het in de sport vaak *common sense* om prestaties, vaak geleverd onder uiteenlopende omstandigheden, te rangschikken. Iedere sport heeft zijn eigen ranglijsten, niet alleen voor huidige generaties sporters, maar ook voor generaties door de jaren heen. Voor beide soorten prestatievergelijkingen worden in dit proefschrift nieuwe technieken gepresenteerd.

In het eerste hoofdstuk worden prestaties van schaatsers tijdens internationale wedstrijden door de jaren heen vergeleken. Sinds 1892 zijn schaatsers jaarlijks actief in internationale toernooien en zijn alle tijden van deze toernooien geregistreerd. Natuurlijk zijn schaatstijden van honderd jaar geleden niet zo maar te vergelijken met huidige prestaties. Overdekte ijsbanen met kunstijs in hooggelegen gebieden hebben, samen met de invoering van de strakke pakken, de klapschaats en de verbeterde trainingstechnieken ervoor gezorgd dat de huidige tijden veel sneller zijn dan de eerste wereldrecords in 1892. Een rangschikking van absolute tijden, zoals in de Adelskalender wordt toegepast, doet dan ook geen recht aan de prestaties van de oude kampioenen.

In het eerste deel van Hoofdstuk 1 worden de invloeden van bovenvermelde veranderingen op de schaatstijden besproken en wordt een maatstaf gedefinieerd, die prestaties door de jaren heen van alle schaatsers vergelijkbaar maakt. Een belangrijk onderdeel van deze maatstaf is dat, voor elke afstand en voor elk toernooi, in plaats van absolute tijden relatieve tijden genomen worden, namelijk de verschillen met de gemiddelde tijd van de beste vijf schaatsers van de betreffende afstand van dat toernooi. Deze verschiltijd, de AV5-waarde genoemd, wordt daarnaast gecorrigeerd voor de algemene ontwikkeling van de sport, het zogenaamde Gould-effect. Globaal gesproken betreft het Gould-effect de afname van onderlinge prestatieverschillen aan de top ten gevolge van een intrinsieke verhoging van het algemene prestatieniveau van de sport.

In het tweede deel van Hoofdstuk 1 wordt besproken hoe AV5-waarden worden gebruikt voor het opstellen van *all time* ranglijsten. Daarbij worden schaatsers beoordeeld op basis van hun beste vier seizoenen, waarbij de resultaten van de belangrijke toernooien, zoals Olympische Spelen, zwaarder worden meegeteld dan bijvoorbeeld World Cups. De beste mannelijke schaatser allertijden is Erik Heiden; bij de vrouwen voert Gunda Niemann de lijst aan. Naast de overal-lijsten zijn ook voor de afzonderlijke afstanden ranglijsten samengesteld.

In Hoofdstuk 2 worden de prestaties van de Nederlandse voetbalinternationals vergeleken. Sinds de invoering van het betaalde voetbal in 1954 zijn de uitslagen van alle internationale wedstrijden op nationaal- en clubniveau, waarin Nederlandse voetballers actief waren, bijgehouden. Op basis van de uitslagen en de speeltijd is bepaald wie de meest succesvolle Nederlandse voetballer is. Omdat voetbal een teamsport is, is het lastig om de individuele bijdrage aan het teamresultaat te bepalen. In dit onderzoek is ervoor gekozen om naar rato van het aantal speelminuten de teambijdrage te bepalen.

De prestaties van de verschillende toernooien worden, net als bij het schaatsen, naar belangrijkheid meegewogen in de eindscore. Zo weegt het spelen van een finale van een Wereldkampioenschap zwaarder dan een finale van een Europa Cup III. Daarnaast zijn structuurwijzigingen van toernooien verwerkt in de scoreverdeling, en wel in die zin dat elk toernooi over de jaren heen hetzelfde aantal punten geeft in de eindscore.

De meest succesvolle Nederlandse voetballer is Ruud Krol, die op de voet wordt gevolgd door Edwin van der Sar. Johan Cruijff bezet de derde plek, maar is wel de best geklasseerde Oranje-aanvaller.

In Hoofdstuk 3 wordt een prestatievergelijkingstechniek gepresenteerd, die tot doel heeft de KNSB te ondersteunen in het selectieproces voor de Olympische Winterspelen. Dit selectieproces is lastig vanwege ondermeer het feit dat er achttien startplekken zijn te verdelen voor vijf afstanden met in totaal slechts tien deelnemers. Dit geldt zowel voor de mannen als de vrouwen. De beperking van tien schaatsers heeft tot gevolg dat de KNSB niet per afstand kan selecteren maar prestaties op verschillende afstanden moet vergelijken.

Voor het selecteren binnen afstanden wordt het Olympisch selectietoernooigebruikt dat elke vier jaar in december gehouden wordt. Maar hoe kies je tussen een 500m-schaatser en een 10000m-schaatser?

In dit hoofdstuk is beschreven hoe op basis van verwachte winstkansen geselecteerd kan worden. De winstkansen zijn berekend uit de uitslagen van op dat moment recente wedstrijden. Vervolgens hebben we selecties (voor mannen en vrouwen) bepaald, waarbij per selectie de totale winstkans op (gouden) medailles maximaal is. Daarnaast hebben we beschreven hoe de KNSB de winstkansen in 2009 heeft gebruikt voor de zogenaamde aanwijslijsten (prestatiematrices in de media). In een aanwijslijst zijn de achttien startplekken van alle afstanden gerangschikt op basis van de berekende winstkansen. Onze 'optimale' selectie week slechts op enkele posities af van de uiteindelijke KNSB-selectie en dit resultaat heeft ertoe geleid dat het model in 2014 opnieuw door de KNSB zal worden toegepast.

In Hoofdstuk 4 wordt naar de eerlijkheid van de 1000m bij het schaatsen gekeken. De beide sprintafstanden, namelijk de 500m en de 1000m, worden tijdens wereldkampioenschappen sprint twee keer per schaatser verreden. Tot 1994 werden op de Olympische Spelen beide afstanden maar een keer verreden. In 1994 heeft de ISU, op basis van het onderzoek van Hjort (2004), besloten de 500m twee keer te laten rijden. In Hjort (2004) is aangetoond dat schaatsers die in de binnenbaan starten een significant voordeel hebben ten opzichte van de buitenbaanstarters. De 1000m wordt echter nog steeds maar één keer verreden tijdens Olympische Spelen. In tegenstelling tot de 500m waarin beide schaatsers een binnenbocht en een buitenbocht rijden, rijden schaatsers tijdens een rit op de 1000m verschillende parcoursen. Binnenbaanstarters schaatsen namelijk een binnenbocht meer.

Met behulp van een regressiemodel is bepaald of dit verschil ook een tijdsvoordeel oplevert. Alle 1000 meters, die twee keer in een weekend zijn gereden tijdens World Cups en Wereld Sprint Kampioenschappen in de periode 2000-2009 zijn gebruikt om te kijken of het starten in de binnenbocht een significant tijdsvoordeel oplevert. Voor de mannen is een niet-significant voordeel van 0.03 seconden gevonden. Bij de vrouwen is het voordeel veel groter, namelijk 0.12 seconden en significant. Naast het feit dat binnenbaanstarters een werkelijk ander traject schaatsen dan buitenbaanstarters, is er met dit gemiddelde tijdsverschil sprake van oneerlijkheid op de 1000m schaatsen. Dit steekt des te meer omdat meer en meer de onderlinge tijdsverschillen op de korte afstanden binnen de foutenmarges van de meetsystemen liggen (ondermeer Gould's effect). Omdat die foutenmarges tenminste 0,003 seconde zijn, zijn tijden die onderling minder dan 0,006 seconde verschillen, feitelijk niet als verschillend aan te merken.